2017

Monitoring Crystal Structure Refinements Using Solid-State NMR Chemical Shift Tensors

Keyton Kalakewich
University of Central Florida

MONITORING CRYSTAL STRUCTURE REFINEMENTS USING SOLID-STATE NMR CHEMICAL SHIFT TENSORS

by

KEYTON J. KALAKEWICH
B.A. Washington & Jefferson College, 2012
M.S. University of Central Florida, 2015

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Chemistry
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Fall Term
2017

Major Professor: James K. Harper

ABSTRACT

Inclusion of lattice-fields in density functional theory (DFT) methods has enabled the accurate calculation of solid-state nuclear magnetic resonance (SSNMR) chemical shift tensors. Calculated ^{13}C and ^{15}N tensors (i.e. 3 principle values per nucleus) can be used to monitor crystal structure refinements and to select the correct structure from a large population of computationally generated candidates. In this dissertation, chapter 2 describes a methodology to improve established crystal structures from three different diffraction techniques involving geometric refinement monitored using SSNMR tensor values. The calculated ^{13}C tensors for three relatively simple organic compounds (i.e. acetaminophen, naphthalene, and adenosine) are shown to markedly improve upon DFT refinement. The so-called GGA-PBE functional provided the best agreement with experimental data. The use of the three principle values of the tensor is required for such results as the average (i.e. the isotropic) is less accurate. Chapter 3 applies this method to differentiate between hundreds of computationally predicted crystal structures. Typically, lattice energy of each candidate is used to select the correct structure, a process which is seldom successful. Herein, it is demonstrated that when ^{13}C tensors from DFT refined structures are used for structural ranking by comparison to experimental data, only the correct structure agrees with experimental data in all cases. Chapter 4 illustrates the use of ^{15}N tensors to monitor DFT refinement as an alternative to the ^{13}C approach of Chapter 2. ^{15}N tensors have been very difficult to obtain previously, thus a novel experimental method is developed here which improves signal-to-noise by as much as 300% and allows routine

measurement. This improvement also improves the accuracy of the tensor values. Overall, the ^{15}N tensors are found to be at least 5 times more sensitive to DFT refinements than ^{13}C values.

ACKNOWLEDGMENTS

I would like to give my thanks to my advisor, Dr. James K. Harper. His tireless support of this work and development of my knowledge as a NMR spectroscopist are two of the main reasons I made it this far. I would like to thank the rest of my committee members: Dr. Seth Elsheimer, Dr. Andres Campigilia, Dr. Artem Masunov, Dr. Karin Chumbimuni-Torres, and Dr. Sean Moore for taking the time to help me improve my written and verbal delivery of both my candidacy and dissertation.

A special thanks goes to Dr. Robbie Iuliucci at Washington & Jefferson College for encouraging me to pursue graduate school at UCF and to work with Dr. Harper. He got me started into this area of research and I certainly wouldn't be here without that initial push. Thank you to STOKES Advanced Research Computing Center here at UCF for partnering with our lab to allow me to run the high level calculations for this and other projects. This work was supported, in part, by the National Science Foundation under grant No. CHE-1455159.

I would also like to thank my fellow group members collectively. I would like to especially recognize Dr. Jacob Powell for his assistance in many joint research projects.

Finally, I would like to thank my wife Becky, for never wavering in her support of my initiative to get my doctorate. Without her, there is no way I would have made it this far.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

CASTEP	Cambridge Serial Total Energy Package
CCDC	Cambridge Crystallographic Data Centre
CIF	Crystallographic Information File
CP/MAS	Cross Polarization – Magic Angle Spinning
CSA	Chemical Shift Anisotropy
FIREMAT	Five Pi Rearranged Magic Angle Turning
GIPAW	Gauge-including Projector Augmented Wave
NMR	Nuclear Magnetic Resonance
PXRD	Powder X-ray Diffraction
RMS	Root Mean-squared
RMSD	Root Mean-squared Deviation
SCND	Single-crystal Neutron Diffraction
SCXRD	Single-crystal X-ray Diffraction
SPINAL	Small-pulse Incremental Alteration
SSNMR	Solid-State Nuclear Magnetic Resonance
TPPM	Two-pulse Phase Modulation

CHAPTER 1: INTRODUCTION

Overview of Dissertation

Lattice fields can strongly influence measurable NMR parameters and it is now possible to include these fields in computed NMR values.[1] The focus of the following chapters in on the methodology required to accurately model SSNMR chemical shift tensors of ^{13}C and ^{15}N nuclei in a variety of molecules. These molecules show a wide range of functional groups, mass, and other structural characteristics. This work has been made possible due to advances in the ability to model the behavior of solid materials computationally by including the effects of the lattice environment. A central point to all the following work is the use of the CASTEP module in Materials Studio to predict tensor values for nuclei of interest and the comparison of these values to experimental values measured in our laboratory. CASTEP employs plane-wave methods to simulate an infinite crystalline lattice in three dimensions and only requires the coordinates of the asymmetric unit contents from a crystallographic information file (CIF). This has the advantage of being both an accessible and simple way to include the lattice-effects of a particular solid. As previously stated, the most important figure of merit in the following studies is the chemical shift tensor. As the following chapters will show, the chemical shift tensor proves to be extremely sensitive to lattice constraints and thus serves as a viable figure to not only improve existing crystal structures but to select the correct one from an array of possible structures. It would be useful to now define the chemical shift tensor and describe how it is measured.

Chemical Shift Tensor

In solid-state nuclear magnetic resonance (SSNMR), it has become common practice to work with powdered samples rather than large single crystals due to difficulties in crystallizing many compounds. This has led to the process of Magic Angle Spinning (MAS), which results in each NMR active nucleus displaying only one line, known as the isotropic line. It is by virtue of different local electronic environments that we can differentiate nuclei of the same chemical identity because the electrons provide a certain amount of shielding from the applied magnetic field. This shielding causes a shift in the signal observed to higher or lower frequencies creating a series of peaks. This shift is referred to as the isotropic chemical shift and, ideally, provides a unique shift for each atomic site in the molecule. This phenomenon is governed by equation 1-1.[2]

$$\mathcal{H} = \gamma_n \frac{h}{2\pi} B_0 I_{z,n}(\boldsymbol{\sigma} - 1) \qquad (1-1)$$

The shielding, $\boldsymbol{\sigma}$, represents the shielding effect of the electrons on the nucleus and is inversely related to the chemical shift, $\boldsymbol{\delta}$. The shielding differs for different orientations in space (i.e. is anisotropic) and is represented by the 2nd rank tensor as shown in equation 1-2.[3]

$$\begin{bmatrix} \boldsymbol{\delta}_{xx} & \boldsymbol{\delta}_{xy} & \boldsymbol{\delta}_{xz} \\ \boldsymbol{\delta}_{yx} & \boldsymbol{\delta}_{yy} & \boldsymbol{\delta}_{yz} \\ \boldsymbol{\delta}_{zx} & \boldsymbol{\delta}_{zy} & \boldsymbol{\delta}_{zz} \end{bmatrix} \xrightarrow{Diagonalize} \begin{bmatrix} \boldsymbol{\delta}_{xx} & 0 & 0 \\ 0 & \boldsymbol{\delta}_{yy} & 0 \\ 0 & 0 & \boldsymbol{\delta}_{zz} \end{bmatrix} \qquad (1-2)$$

This matrix represents components of the shielding (or shift) arising from a nuclear spin in a particular orientation in the applied magnetic field, B_0. While working with the complete tensor is desirable, the work here focuses on the use of the diagonal elements in the matrix known as the principle values (i.e. δ_{xx}, δ_{yy}, δ_{zz}). Mathematically, diagonalizing the complete

tensor results in the loss of orientational information, thus, the Cartesian designation (δ_{xx}, etc.) becomes ambiguous and it is more convenient to organize the principle values by magnitude such that $\delta_{11} \geq \delta_{22} \geq \delta_{33}$. Differences between these principle values are due to differences in the bonding around a given nucleus and is thus called the chemical shift anisotropy (CSA). These tensor components have sufficient sensitivity to accomplish the aims of the studies herein and have the advantage of being easy to acquire in NMR experiments, unlike the off-diagonal terms, which come from single crystal NMR experiments. Spectroscopically, these quantities may be acquired from a variety of methods with the simplest being a single-pulse cross-polarization/magic-angle spinning (CP/MAS) experiment. This method generates a spectrum whose appearance changes with the spinning rate of the sample as shown in figure 1-1.

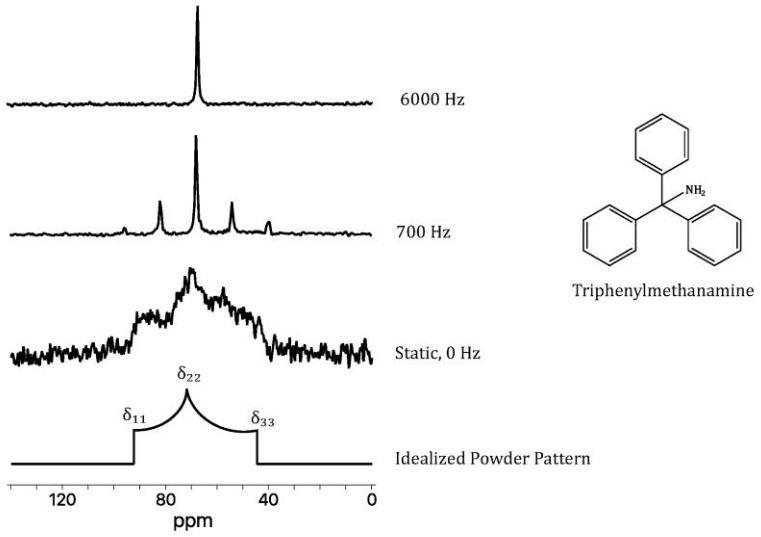


Figure 1-1: An illustration of the effect of spinning speed on the ^{15}N spectrum of Triphenylmethanamine (also shown to the right). As the spinning speed decreases, the single isotropic signal devolves into a spinning side band pattern with the lines separated by the spinning speed in Hz until, under static conditions, the powder pattern is produced. An idealized version is shown with labels for the three principle components.

The individual lines in figure 1-1 are referred to as "spinning-sidebands" and are separated by the spinning speed in Hz. Many SSNMR experiments focus only on the isotropic shift (i.e. $1/3(\delta_{11}+ \delta_{22}+ \delta_{33})$), attained by spinning quickly (speeds of several kHz or more) at the "magic angle" of 54.7°. This angle arises from orienting the spinning axis inside a cube with the axis running through opposite corners of the cube and its center.[4] This allows each orthogonal axis to be rotated into one another during a rotor period resulting in the principle components being averaged to give the isotropic chemical shift. As the rotor speed decreases more sidebands become resolved. Once the spinning speed reaches zero (static conditions), a static powder pattern corresponding to a series of overlapping sidebands is obtained. From this static spectrum, the principle values can be directly obtained from the positions labeled in figure 1-1. While the acquisition of static spectra would be ideal, it is not practical in many cases because the area of the static pattern is conserved thus spinning improves signal-to-noise (s/n) by concentrating the signal into fewer resonances. Thus, slow spinning experiments in conjunction with sideband fitting software, such as HBA 1.6,[5] can simulate the spinning sideband pattern and calculate the three principle values for a nucleus from sideband intensities while allowing shorter experiment times than the same compound under static conditions.

Chemical Shift Anisotropy and Crystallography

Currently, the most widely used methods of structure determination in solids are single-crystal neutron diffraction, single-crystal x-ray diffraction, and powder x-ray diffraction (SCND, SCXRD, and PXRD, respectively). These diffraction methods determine the position of atoms in a crystal lattice. While these methods are mature and rapid, they are not entirely without deficiencies. For example, SCXRD typically provides relatively inaccurate hydrogen positions. Modern PXRD is relatively inaccurate in all atomic positions and usually makes no attempt to locate hydrogens. In contrast, SCND provides accurate coordinates for all atomic sites and is considered the "gold-standard" for structural analysis. Unfortunately, neutron sources are rare and thus SCND structures comprise less than 0.5% of the structures in the Cambridge Crystallographic Data Centre (CCDC). This limitation makes diffraction data of variable quality and results in many structures in the CCDC that could benefit from improved methods. The use of SSNMR tensors and CSA to monitor the accuracy of a crystal structure provides much greater sensitivity than most diffraction methods and has the advantage of being easy to acquire and calculate.

SSNMR Tensor Acquisition and Calculation

Currently, there are a remarkable variety of SSNMR methods and each can yield a different type of information. Each pulse sequence is designed to accomplish a specific goal and range from simple one-dimensional acquisitions to sequences such as the five-pi rearranged magic-angle turning (FIREMAT) sequence (used herein), which contains multiple pulses and dimensions.[6]

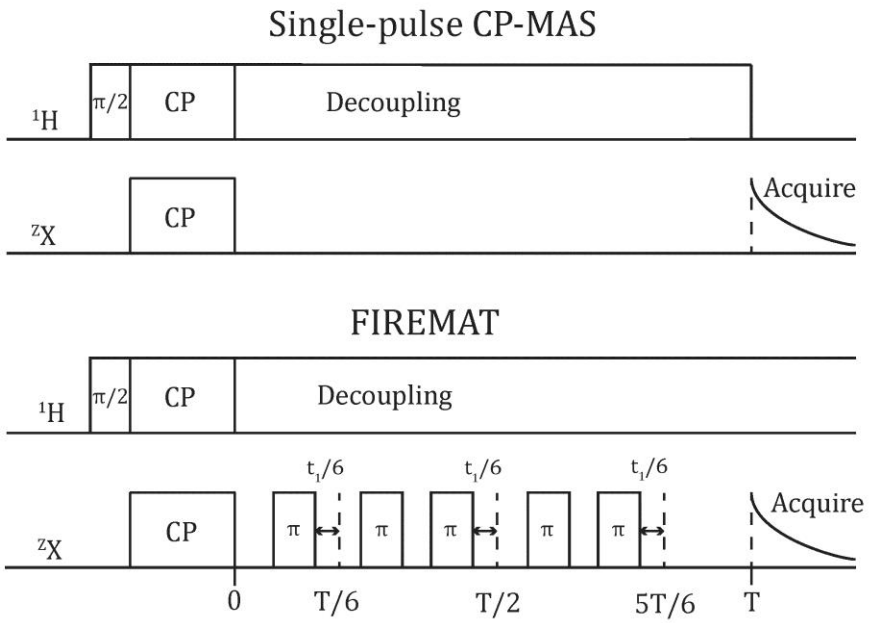


Figure 1-2: Pulse sequence diagrams for single-pulse CP-MAS and FIREMAT sequences are shown. Pulse angles are shown in radians rather than degrees in the rectangles. T represents the rotor period and t_1 represents the evolution time. Decoupling may be either TPPM or SPINAL

Figure 1-2 illustrates the difference between these two sequences. The advantage of using CSA as a figure of merit in crystallographic studies is that it can be acquired using the single pulse CP/MAS experiments in which magnetization is transferred from an abundant, NMR-active nucleus (^{1}H) to less abundant or insensitive nucleus (e.g. ^{13}C or ^{15}N, respectively).

This allows us to measure the chemical shift of these nuclei much more quickly than we could do directly. In addition, slow-spinning or static conditions provide the CSA with 3 parameters for each NMR-active nucleus. The use of a single-pulse CP/MAS experiment is effective for molecules with only a few resonances, but is less useful when there are many resonances in the spectrum. In such spectra, the powder or sideband patterns can overlap and thus may require more complex pulse sequences to properly assign sites. For such complex scenarios, FIREMAT is a preferred sequence. It is a two-dimensional experiment, which artificially creates the second dimension using a modified pseudo 2D spinning sideband (P2DSS) rearrangement method coupled with the technique for importing greater evolution resolution (TIGER).[7,8] This combination results in a reduced amount of data collection while extending the indirect dimension within the same amount of time, T. This sequence, while useful for heavier molecules with multiple sites, requires more complex and careful calibration because with each additional pulse any error in calibration will propagate through the experiment. Regardless of how the experimental data are acquired, it provides a benchmark that can be used as the "true" value for the tensor of one or more nuclei. These data thus guide the development of computational approaches for calculating tensors and for the use of these tensors in structure refinement and selection.

In general, geometry optimizations will be governed by parameters such as distance, forces on atoms, and overall energy decreases to make changes to a structure iteratively until a minimum is reached. As an illustration, Figure 1-3 shows cimetidine before and after a typical geometry optimization. The changes in atomic positions are larger than diffraction errors for hydrogens as well as heavy atoms. After optimization, the

SSNMR parameters can be calculated using the software, CASTEP, in Materials Studio.[9] These calculated tensors can be compared to experimental values to determine the viability of the calculated model. There are many adjustable constraints to each of these calculations including the choice of functional and energy cut-offs this flexibility provide a wide range of options depending on the structure being assessed. In the following chapters, it was found that the PBE functional and the "ultra-fine" level of calculation provided structural solutions that matched experimental data of both ^{13}C and ^{15}N with errors of 3.0 ppm and 7.0 ppm, respectively. When the published crystal structure was compared to experimental data without first being optimized, the ^{13}C and ^{15}N errors increased by roughly 10.0 and 47.0ppm, respectively, thus illustrating the need for the geometry optimization step.[10,11,12] NMR-RMS differences was only one figure of merit used to monitor the efficacy of this method while the use of force on atoms and individual displacements help to corroborate the improvement in the structures. Some of those displacements were significant as shown in Figure 1-3. It is noteworthy that ^{15}N tensors show a vastly superior sensitivity to ^{13}C tensors, with average changes of 47.0ppm for ^{15}N versus 8.0ppm for ^{13}C. This differential is presumed to be due to the higher range of CSA values for nitrogen compared with that of carbon, which comes largely from the presence of a lone pair of electrons present on neutral sp^3-hybridized nitrogens.

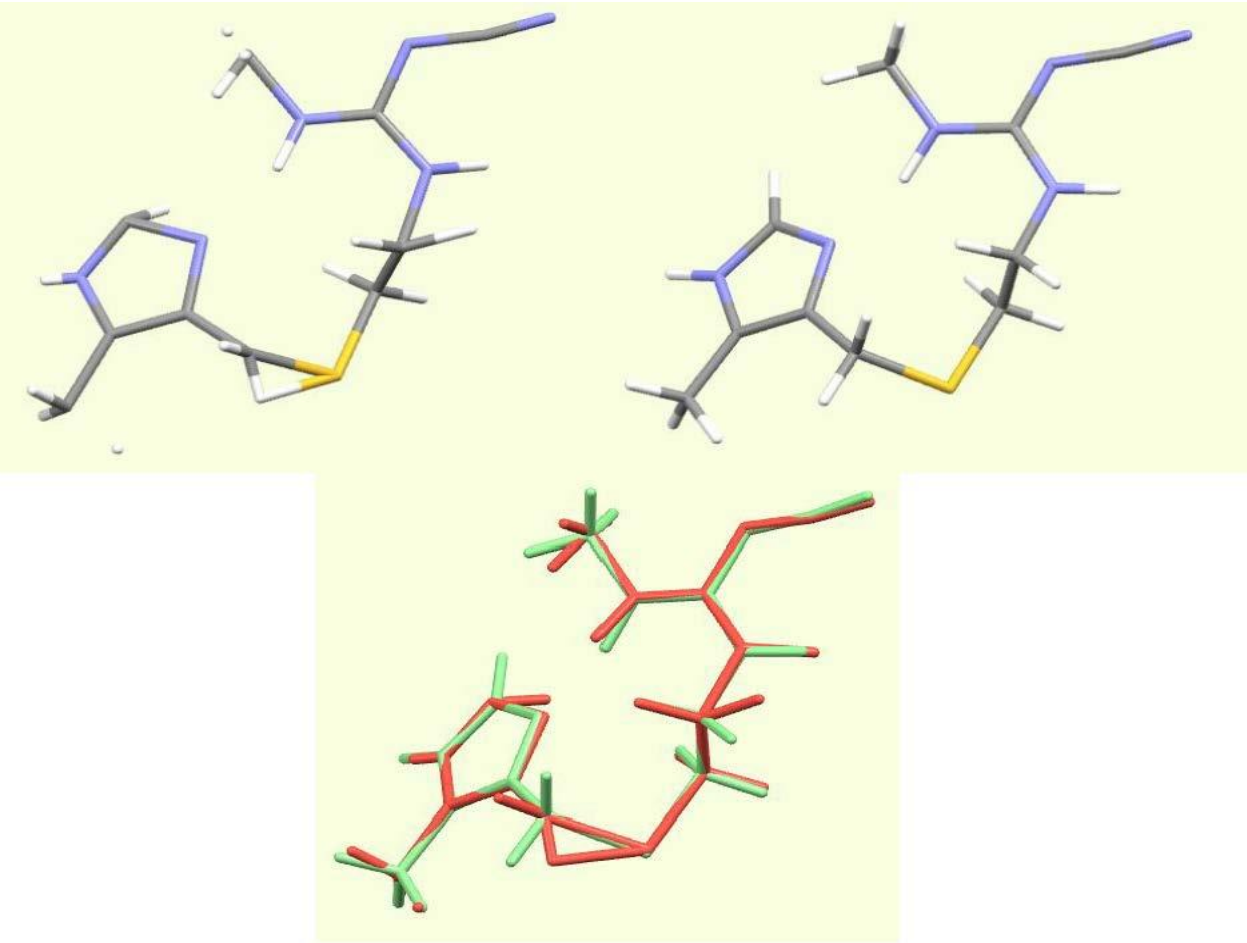


Figure 1-3: The structures of cimetidine before and after geometry optimization are show as the upper left and upper right structures, respectively. The lower portion represents a structural overlay of the "before" and "after" structures. Red is before optimization and green is after.

Chapters Appearing as Publications

The remaining chapters appear as they were originally published in various peer review journals. Chapters 2, 3, and 4, were published, respectively, in CrystEngComm, Crystal Growth & Design, and Physical Chemistry Chemical Physics.[10, 11, 12]

References

[1] Pickard, C. J.; Mauri, F. *Phys. Rev. B: Condens. Matter Mater. Phys.* 2001, 63, 245101.
[2] Keeler, J. *Understanding NMR Spectroscopy,* 2nd ed. Wiley, Hoboken, 2013.
[3] Alderman, D. W.; Sherwood, M. H.; Grant, D. M. *J. Magn. Reson.* 1993, 101, 188.
[4] Bax, A.; Szeverenyi, N. M.; Maciel, G. E.; *J. Magn. Reson.* 1983, 52, 147.
[5] HBA 1.6, K. Eichele, Universität Tübingen, 2015.
[6] Alderman, D. W.; McGeorge, G.; Hu, J. Z.; Pugmire, R. J.; Grant, D. W. *Mol. Phys.* 1998, 95, 1113.
[7] Gan, Z. *J. Mag. Reson.* 1994, 109, 253.
[8] McGeorge, G.; Hu, J. Z.; Mayne, C. L.; Alderman, D. W.; Pugmire, R. J.; Grant, D. W. *J. Magn. Reson. 1997, 124, 134.*
[9] Accelrys Software Inc., Materials Studio, Release 6.0, San Diego, 2011.
[10] Harper, J. K.; Iuliucci, R.; Gruber, M.; Kalakewich, K. *CrystEngComm.* 2013, 15, 8693.
[11] Kalakewich, K.; Iuliucci, R.; Harper, J. K. *Cryst. Growth Des.* 2013, 13, 5391.
[12] Kalakewich, K.; Iuliucci, R.; Mueller, K. T.; Eloranta, H.; Harper, J. K. *J. Chem. Phys.* 2015, 143, 194702.

CHAPTER 2: REFINING CRYSTAL STRUCTURES WITH EXPERIMENTAL ^{13}C NMR SHIFT TENSORS AND LATTICE-INCLUDING ELECTRONIC STRUCTURE METHODS

Reproduced from Harper, J. K. et al. Refining crystal structures with experimental ^{13}C NMR shift tensors and lattice-including electronic structure methods. *CrystEngComm*, 2013, 15, 8693. With permission from the Royal Society of Chemistry

James K. Harper,[a,*] Robbie Iuliucci,[b] Matthew Gruber,[b] Keyton Kalakewich[b]

[a]*The University of Central Florida, Department of Chemistry, Orlando FL 32816 and* [b]*Washington and Jefferson College, Department of Chemistry, Washington, PA 15301.*

Abstract

Large differences are found in the quality of crystal structures obtained from different diffraction methods. The most accurate studies identify all atomic sites, including hydrogens, while others lack even the resolution needed to locate individual atoms. The gauge including projector augmented wave method (GIPAW) provides a technique to further refine any of these structures under lattice constraints. Here, the sensitivity of solid-state NMR ^{13}C shift tensor principal value data to GIPAW refinement is investigated. The refinement is shown to improve x-ray powder, x-ray single crystal and even neutron single crystal diffraction data. Convergence to a single structure is observed in most cases. Surprisingly, the final refined structures usually diverge from the original neutron diffraction coordinates - data typically viewed as the most accurate. To ensure that the structural changes represent improvements, three metrics are monitored comprising fit to ^{13}C shift tensors, forces upon the atoms and changes in atomic positions relative to a reference structure. In all cases these parameters improve upon refinement suggesting

that GIPAW creates structures surpassing the accuracy of single crystal neutron diffraction data. However, the influence of thermal motions remains unknown. Improvements are seen most strongly in forces and NMR fits and least in atom positions. This study evaluates reasonably accurate model structures to quantify improvements. However, structures obtained from lower resolution methods (*e.g.* electron diffraction) will benefit most from GIPAW refinement. In such structures the refinement has the potential to convert structures with questionable atom positions into coordinates rivaling neutron diffraction single crystal data.

Introduction

Since the first crystal structures from diffraction were reported in 1913,[1] crystallography has grown into a discipline with an extraordinarily far reach and a powerful influence on the development of modern science. Insights gained from diffraction studies have provided some of the most important guiding principles in solid-state physics, biology, and earth science. In chemistry, crystallography provided the first experimental distinction between a covalent and an ionic bond,[2] established values for atomic and ionic radii,[3] and provided the evidence needed to unambiguously assign absolute stereochemistry.[4] Despite crystallography's long history, it remains a discipline that can produce results of fundamental importance. For example, recent work has demonstrated the existence of a new type of unusually long carbon-carbon bond in molecule-based magnets.[5] Likewise, the existence of C – H ⋯ X hydrogen bonds (X = O, N, and Cl) was only firmly established in 1982 based on extensive surveys of crystallographic databases.[6] Most intriguing, perhaps, is the discovery of quasicrystals, materials that challenge the fundamental concept of what constitutes a crystal.[7] Considering the powerful insights provided by crystallography, it is not surprising that diffraction studies have become the method of choice for many seeking information on chemical structure and function.

Many of the advantages of crystallography come from its ability to provide structural information on individual atomic positions. In contrast, most other analytical methods identify only functional groups or provide information on certain properties of a material. Currently, crystal structures are obtained by several methods including single

crystal diffraction and powder (or fiber) diffraction using x-rays, neutrons or electrons. In some cases, solid-state NMR (SSNMR) can also provide crystal structure and a few structures have now been derived entirely from SSNMR data.[8] SSNMR shares with diffraction methods the ability to acquire information on individual atomic sites. Another technique for obtaining crystal structures is a theoretical approach known as crystal structure prediction (CSP).[9] At present, CSP methods are capable of making predictions in moderate molecular weight compounds (*i.e.* < 1 kDa) with some programs capable of generating structures in all possible space groups. However, selection of a single structure matching an experimentally observed crystal remains a challenge and further development is needed.[9] In general, the structures obtained from powder diffraction, SSNMR data and CSP techniques are all of lower quality than those derived from single crystal diffraction data. Certain single crystal structures are also of poor quality. For example, single crystal data for most proteins have lower resolution than corresponding data for small molecules, yet much of this deficiency is difficult to eliminate by conventional diffraction techniques. Likewise, many structures in small molecule single crystal databases have large R-values, indicating the potential for improvement. Thus, a wide variety of crystal structures could benefit from further refinement and the development of complementary methods for improving these structures is an important endeavor.

Considerable work suggests that SSNMR data provides a very accurate reflection of crystal structure and it has been demonstrated that including SSNMR structural constraints in a refinement can improve structure.[10] In early studies, SSNMR data provided information on certain structural features that were difficult to establish by

crystallography. Examples include measurement of selected distances (*e.g.* $^{129}Xe/^{129}Xe$ & $^{29}Si/^{19}F$),[11] detection of static disorder in a small region of an otherwise ordered material,[12] determination of the tautomeric form of a molecule,[13] and identification of the number of molecules in the asymmetric unit.[14] These studies facilitated further refinement of existing crystal structures.

Recently an approach to crystal structure refinement has been developed that allows for more comprehensive improvements. Initial efforts involved creation of a variety of structural candidates and then assessed each based on the agreement between experimental SSNMR parameters and the same parameters computed from the model structure.[15] While these approaches are accurate, they rely on software that does not consider the lattice and require input from an analyst capable of identifying probable conformations. Recently, it has become possible to refine entire structures using efficient methods that include lattice effects[16,17] and this approach has become widely accepted. The method most commonly employed is known as the gauge including projector augmented wave (GIPAW) and requires only an initial crystal structure. GIPAW refinements have been demonstrated to improve structures derived from powder diffraction,[18] SSNMR[19] and even single crystal neutron diffraction data.[20] In most cases the changes in the crystal structure are small and usually create almost no observable change in the x-ray powder diffraction pattern.[18a,18b,21] However, these refinements usually significantly improve the fit between computed NMR parameters and experimental data. Taulelle has asserted that, in general, powder diffraction data will be less sensitive than

SSNMR data[22] and many of the early studies on structural refinement supports this conclusion.

Currently, the types of SSNMR data that have been used to monitor refinement of crystal structures include isotropic shifts (^{27}Al, ^{29}Si, ^{13}C & ^{31}P),[16,18a,23] ^{1}H/^{1}H spin diffusion,[19] quadrupolar coupling constants (^{17}O & ^{27}Al),[16,18a] pseudocontact shifts (^{59}Co),[24] dipole coupling data (^{29}Si/^{29}Si)[15b] and the chemical shift tensor using both principal values (^{29}Si)[18c] and the full tensor (^{13}C)[20] consisting of 6 shift parameters per nucleus. It is interesting to note that certain SSNMR parameters are more sensitive to the GIPAW refinement than others. For example, Brouwer directly compared a structure derived from ^{29}Si/^{29}Si dipole-coupling data with the same structure refined using ^{29}Si tensor principal values and found that principal values gave a more accurate structure.[15b] Another study involving ^{1}H data found that isotropic chemical shifts were more sensitive to structural refinement than ^{1}H/^{1}H spin diffusion data.[19] In general, it is also well known that tensor principal value data are more sensitive than isotropic shifts in establishing structure. This sensitivity of tensor data to refinement was superbly illustrated by the work of Johnston et al.[20] where the agreement between experimental and computed full ^{13}C shift tensor data (*i.e.* 6 shifts per nucleus) improved significantly when a GIPAW refinement was performed. Remarkably, the initial structures in this study were high quality single crystal neutron diffraction structures, suggesting that further refinement of even very accurate crystal structures may be possible if appropriate SSNMR data are employed. This ^{13}C tensor study is particularly compelling because it includes an extensive data set composed of *all* ^{13}C sites in 14 different compounds. Hence it appears that the structural improvements observed

are not anomalous observations due to analysis of an unusual compound, but results that can be generally expected. Unfortunately, the full tensor data is rather difficult to obtain experimentally as it usually requires large single crystals and specialized equipment. Here we wish to explore the sensitivity of a more easily acquired set of tensor values to GIPAW refinement, namely the ^{13}C tensor principal values consisting of three shifts per nucleus. Our aim in this study is to determine if the sensitivity to GIPAW refinement observed with the full tensor data is retained when ^{13}C tensor principal value data are used. Brouwer's prior work on ^{29}Si tensor principal values in zeolites has already demonstrated that such data are highly sensitive to refinement and his study incorporating principal values produced a structure of comparable quality to that obtained from single crystal x-ray diffraction. We anticipate that a similar sensitivity to refinement will be observed with ^{13}C tensor data and that such analyses will extend these refinement methods to a wider range of compounds than can be evaluated by ^{29}Si tensor data. To ensure that improved agreement between computed and experimental ^{13}C principal values reflects genuine improvements in lattice structure, other parameters including forces on atoms and average changes in atomic positions in the lattice will also be monitored.

In the present study, an analysis of the sensitivity of ^{13}C principal values tensor data to GIPAW refinement is performed on three model compounds that all have well-established single crystal neutron diffraction, single crystal x-ray diffraction and x-ray powder diffraction structures. These refinements establish where improvements can be observed, quantify the changes in computed principal value shifts and other parameters and demonstrate that this process improves all three types of diffraction structures.

Results and Discussion

To evaluate the sensitivity of ^{13}C tensor principal values to GIPAW refinement, naphthalene, acetaminophen and adenosine (see Figure 2-1) were selected as model compounds because each has an established x-ray powder[25,26,27] x-ray single crystal[28,29,30] and neutron single crystal[31,32,33] diffraction structure for the same crystalline phase. Accordingly, these compounds provide the information needed to evaluate the outcome of GIPAW refinement on structures obtained from a variety of diffraction techniques. Of these data types, single crystal neutron diffraction structures are usually regarded as the most accurate. The availability of neutron data therefore allows us to verify that GIPAW refinements result in genuine improvements to structure through comparison of atomic positions in refined structures with corresponding sites in the neutron data.

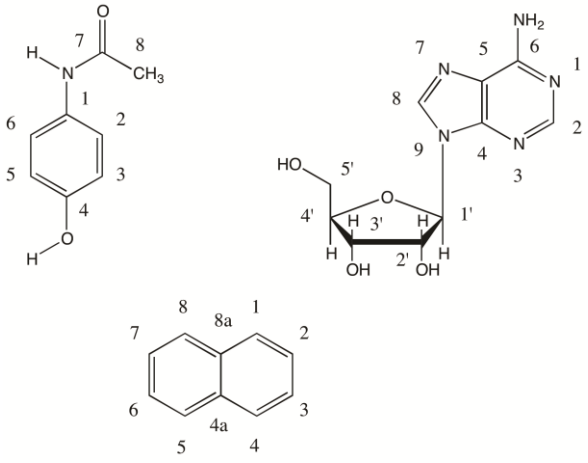


Figure 2-1: Structures of acetaminophen (top left), adenosine (top right) and naphthalene used as model structures to explore the influence of GIPAW refinement on ^{13}C chemical shift tensor principal values. Each structure has a previously established x-ray powder, single crystal x-ray and single crystal neutron diffraction structure for the same phase, allowing study of the influence of refinement on structures obtained from a wide variety of techniques.

Experimental ^{13}C chemical shift tensor data for adenosine and naphthalene have been reported previously using, respectively, FIREMAT[34,35] and a static single crystal NMR method.[36] Experimental errors in the naphthalene values are reported to be ± 0.54 ppm[37] while the errors for FIREMAT data have been reported to be ± 0.9 ppm.[15d] Because both datasets were externally referenced to the TMS scale and collected by the same research group, chemical shift tensor values are presumed to be of comparable accuracy and were used here without modification. The experimental ^{13}C tensor principal values for acetaminophen were also measured here using the FIREMAT experiment (Table 2-1) and externally referenced to the TMS scale.

The possibility of polymorphism complicates the proposed study since care must be taken to ensure that the same phase is used in both the SSNMR and diffraction studies. Here, x-ray powder diffraction data were acquired for model compounds having more than one known polymorph to verify that the form used for SSNMR analysis matched diffraction data (see Experimental).

Accurate structural analyses require correct assignment of shifts to molecular positions. The SSNMR ^{13}C assignments for adenosine and naphthalene are known from prior work.[34,36] For acetaminophen, ^{13}C shift assignments for most sites were obtained from a $^{1}H/^{13}C$ heteronuclear correlation (HETCOR) spectrum acquired using a 50 μs cross-polarization time. All carbons directly bonded to protons gave strong correlations. Seven longer-range interactions were also observed involving the positions illustrated in Figure 2-2. These longer correlations allowed all resonances except C2, C3 and C6 to be assigned. Assignment at C2 was made based on the observation of a $^{1}H \rightarrow C1$ correlation that

restricts the signal at 123.6 ppm to C2 or C6. Computed ^{13}C tensor principal values from C2 matched experimental data significantly better than those from C6, assigning the resonance at 123.6 ppm to C2. Likewise, a $^{1}H \rightarrow$ C4 correlation restricted C3 to either the resonance at 115.7 or 120.6 ppm. A similar comparison between computed and experimental shift tensors assigned the resonance at 115.7 ppm to C3. These assignments left only the resonance at 120.6 unassigned and this peak was assigned by default to C6. Final shift assignments for acetaminophen agree with those previously reported[38] except at C2, C3, C5 and C6 where the shift assignments were reversed (i.e. C2↔C6 and C3↔C5). The new acetaminophen assignments (Table 2-1) are used in all analyses described herein. All ^{13}C shift assignments and tensor principal values are listed in Tables 2-1, 2-2, and 2–3. A HETCOR spectrum is included in Electronic Supplementary Information together with ^{1}H shift assignments.

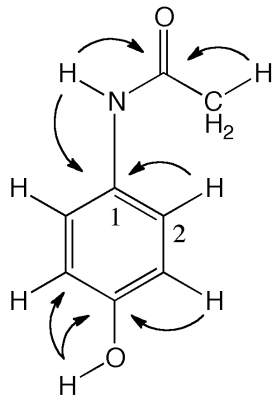


Figure 2-2: The longer-range $^{1}H/^{13}C$ correlations observed in acetaminophen. One-bond $^{1}H \rightarrow ^{13}C$ correlations were also observed for all protonated carbons, but have been omitted for clarity.

Table 2-1: The ^{13}C shift tensor principal values (ppm) for acetaminophen, experimental (theoretical[a]).

Position	Carbon type	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)[b]
C1	Quat.	228.9 (219.0)	133.4 (136.8)	36.8 (35.5)	133.1 (130.4)
C2	CH	203.5 (207.9)	150.4 (152.0)	16.2 (3.7)	123.4 (121.2)
C3	CH	188.8 (194.3)	133.4 (128.3)	24.7 (16.7)	115.7 (113.1)
C4	Quat.	238.1 (231.2)	154.2 (161.5)	64.6 (63.2)	152.3 (152.0)
C5	CH	193.5 (196.0)	130.0 (129.3)	25.4 (13.2)	116.4 (112.8)
C6	CH	194.8 (205.1)	139.4 (135.9)	27.5 (14.8)	120.6 (118.6)
C7	Quat.	245.8 (236.7)	174.8 (171.0)	88.8 (88.3)	169.8 (165.3)
C8	CH_3	46.5 (41.8)	27.5 (27.8)	-2.5 (-4.7)	23.8 (21.7)

[a]Theoretical shifts were computed using neutron diffraction single crystal coordinates after GIPAW refinement. [b]Isotropic shifts at positions 2, 3, 5 and 6 differ from the assignments originally reported by Jagannathan.[38] The shift values listed above were used for all comparisons herein.

Table 2-2: ^{13}C shift tensor principal values (ppm) for adenosine,[a] experimental (theoretical[b]).

Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C2	239 (232.7)	158 (164.3)	66 (58.8)	154.8 (151.9)
C4	221 (214.9)	166 (161.2)	58 (57.8)	148.5 (144.7)
C5	167 (159.4)	145 (153.9)	47 (45.7)	119.7 (119.7)
C6	222 (201.8)	191 (193.7)	52 (52.7)	155.2 (149.4)
C8	216 (212.8)	136 (135.1)	61 (60.7)	137.8 (136.2)
C1’	109 (110.6)	92 (100.5)	76 (74.1)	92.3 (95.1)
C2’	100 (107.3)	75 (74.6)	38 (34.6)	71.2 (72.1)
C3’	88 (89.5)	78 (80.9)	59 (59.1)	75.0 (76.5)
C4’	113 (115.3)	93 (97.5)	48 (45.1)	84.9 (86.0)
C5’	86 (90.6)	68 (66.8)	34 (31.7)	62.7 (63.0)

[a]Shift assignments and tensor data were originally reported by Stueber and Grant.[34] [b]Theoretical shifts were computed using neutron diffraction single crystal coordinates after GIPAW refinement.

Table 2-3: ^{13}C shift tensor principal vales for naphthalene,[a] experimental (theoretical[b]).

Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1, C5	224.7 (224.0)	140.3 (143.5)	22.8 (17.4)	129.3 (128.3)
C2, C6	227.6 (226.8)	139.3 (139.7)	11.1 (3.9)	126.0 (123.5)
C3, C7	227.6 (226.5)	138.2 (139.7)	10.4 (3.5)	125.4 (123.2)
C4, C8	223.9 (222.7)	145.6 (149.1)	20.5 (14.3)	129.9 (128.7)
C4a, C8a	208.5 (205.4)	202.2 (202.0)	-5.9 (-12.5)	134.9 (131.6)

[a]Shift assignments and tensor data were originally reported by Sherwood et al.[36]
[b]Theoretical shifts were computed using neutron diffraction single crystal coordinates after GIPAW refinement.

Refining x-ray powder diffraction structures

All structural refinements were performed using GIPAW as implemented in the CASTEP density functional theory code.[39] The generalized gradient approximation (GGA) PBE[40] method was a logical choice for structural refinement because it is known to yield very accurate ^{13}C tensor data when full tensors are computed.[20] To verify that GGA-PBE produces accurate structures and to compare its performance with other GIPAW methods, refinements using the GGA-PBE, GGA-PW91,[41] and LDA-CA-PZ[42] methods were performed. This process involved refining the powder structures of the three model compounds by each method, then comparing the resulting structure to the neutron diffraction coordinates. The root-mean-square deviations (RMSDs) in atomic coordinates before and after refinement were evaluated. The PBE and PW91 methods gave the largest overall improvements and nearly identical results for all three crystals. The older LDA method produced a significantly worse adenosine structure but comparable coordinates for acetaminophen and naphthalene (Figure 2-3).

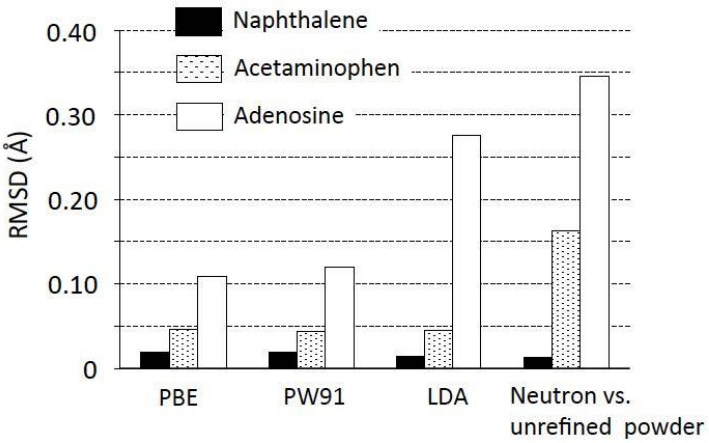


Figure 2-3: A comparison of atomic positions (RMSDs, Å) showing neutron single crystal coordinates versus GIPAW refined powder positions using the PBE, PW91 and LDA methods and unrefined powder data. The RMSD values were obtained by comparing all atoms in the superimposed structures.

To place the RMSDs of the refined powder structures in context, an RMSD of 0.01 – 0.1 Å is typically found when comparing two independently acquired high quality single crystal diffraction structures of the same moderately sized organic compound.[43] Thus the PBE and PW91 refinements have converted these powder structures into structures having coordinates that lie within the error of the single crystal neutron diffraction values.

The refinement of naphthalene is unique among the model compounds in that the RMSD between powder and neutron diffraction coordinates *before* refinement is negligible (i.e. 0.013 Å). In fact, the GIPAW refinement appears to slightly degrade the powder structural quality. However, RMSD is a figure of merit that includes all positions. Since NMR provides information for each atomic site, a more insightful use of the data involves a corresponding analysis of individual atomic positions. Such analysis reveals that, in fact, the C4a–C8a bond in naphthalene lengthens 0.016 Å upon GIPAW refinement. Similarly, the C1–C2 bond (along with the symmetry related C5–C6 bond) increases by 0.010 Å. These adjustments are larger than the estimated error in C–C bond lengths of ± 0.001–0.005 Å and are thus of significance. All other GIPAW adjustments to C–C distances are comparable to the error.

While these adjustments to naphthalene are small, comparison of other figures-of-merit (described hereinafter) suggests that this refinement is a genuine improvement to the structure. This potential improvement to naphthalene's structure is somewhat surprising since napthalene has been the subject of intense study since its structure was first solved in 1949.[44] Currently, 34 diffraction structures have been reported[28,31,44,45] with

2 additional studies on deuteronaphthalene.[46] Further refinement of such a well-studied structure, although minor, demonstrates the sensitivity of this methodology. This analysis emphasizes that RMSD of coordinates is a figure of merit that has limitations and other evidences were sought to establish whether the GIPAW refinement has actually improved the structure.

Another relevant metric in assessing improvement in a crystal structure is forces upon the atoms in the lattice. Reduction of forces is part of CASTEP's geometry optimization algorithm and can therefore be evaluated at any step in a refinement. In the model structures, GIPAW refinement of the powder data was found to decrease forces on the atoms in all cases by two orders of magnitude with PBE showing the largest improvements (see Electronic Supplementary Information). These data are plotted in Figure 2-4 for the PBE refinement, together with SSNMR fit (described below) and a comparison of RMSDs. Ashbrook et al. have used the decrease of these forces as evidence of structural improvement,[18a] but also report a case where a decrease in forces did not correspond with a better fit to computed NMR parameters.[47] Nevertheless, when considered together with improvements in atomic positions relative to a reference structure, these forces provide a useful indicator of structural improvement.

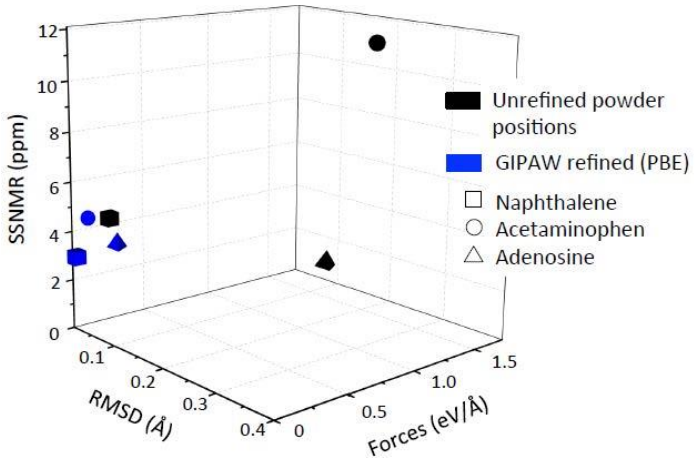


Figure 2-4: A plot showing the decrease in forces on the atoms, RMSDs relative to neutron single crystal positions and improvement in SSNMR fit between experimental and computed ^{13}C tensor data resulting from GIPAW refinement of the powder diffraction coordinates. Only PBE results are shown here. A complete list of data from all methods is included as Electronic Supplementary Information.

A key point in our analysis is verification that the improvements in structure, described above, correspond to better agreement between computed and experimental ^{13}C tensor principal values. This is an important consideration because the GIPAW refinement does not explicitly consider the fit to any type of NMR data in its convergence, thus the final structure is not required to be in agreement with experimental NMR evidence. However, several authors have now demonstrated that the GIPAW refinement usually correlates with improvement in SSNMR parameters.[18,19,20] A comparison of differences between experimental tensor data and values calculated after GIPAW refinement demonstrates that most refinements do, in fact, improve the fit with improvements ranging from 0.20 – 6.21 ppm. The PBE and PW91 functionals usually gave the largest improvements (see Electronic Supplementary Information). With the exception of adenosine, the improved fit is statistically significant and allows refined structures to be distinguished from unrefined coordinates with greater than 90% confidence. These data, taken together with the data on

forces and atom positions already described, establish that *improvement in structure from GIPAW refinement corresponds with better agreement between computed and experimental ^{13}C tensor principal values*. The PBE and PW91 give nearly identical results and both these methods usually give a better fit than the older LDA method. Since PBE is reported to correct some instabilities of the PW91 method,[40] GGA-PBE was selected for all calculations reported herein. Improvements in the SSNMR fit due to PBE refinement are illustrated in Figure 2-4.

It is noteworthy that the fit of the computed ^{13}C tensor data for acetaminophen and adenosine is inferior to that observed for naphthalene. This is attributed, in part, to the fact that experimental principal value data for ^{13}C sites bonded to ^{14}N are known to have larger errors than compounds containing only C, H, and O due to $^{14}N/^{13}C$ dipole coupling.[48] In adenosine, 60% of the ^{13}C sites are bonded to a ^{14}N, while nearly a third of the carbons in acetaminophen have a bonded ^{14}N. In this study, the primary focus is on the *improvement* in the fit between experimental and computed ^{13}C tensor data upon refinement as a means of monitoring crystal structure quality. The absolute magnitude of the improvement is less emphasized.

Refining single crystal x-ray diffraction structures

To better assess the influence of the refinements, an additional evaluation was performed for each x-ray single crystal model structure that involved GIPAW refinement and comparison to the neutron diffraction coordinates. It was anticipated that x-ray single crystal diffraction data would show improvements upon refinement similar to those observed for powders because hydrogen coordinates are often not well determined by x-

ray single crystal diffraction. The single crystal x-ray structures of all model compounds did, in fact, improve based on comparisons of RMSDs of atomic positions, forces and fits to SSNMR data. Figure 2-5 illustrates the improvements in all parameters from refinement. It is somewhat surprising to find that improvements involved not only hydrogen positions, but also many heavy atoms. The contribution of *only* hydrogen positions to the overall refinement was evaluated by performing a separate GIPAW analysis where only hydrogen coordinates were refined (right plot in Figure 2-5). Comparable analyses were also performed on the powder structures to assess where the largest adjustments were occurring in these data. The changes in atomic positions (RMSD) are illustrated in Figure 2-6. These data demonstrate that the largest structural changes in both the single crystal and powder data involve adjustments in hydrogen positions. However, most structures also exhibit significant changes to heavy atom positions. A complete listing of changes in force, atomic positions, and the fit to SSNMR data and is included as Electronic Supplementary Information.

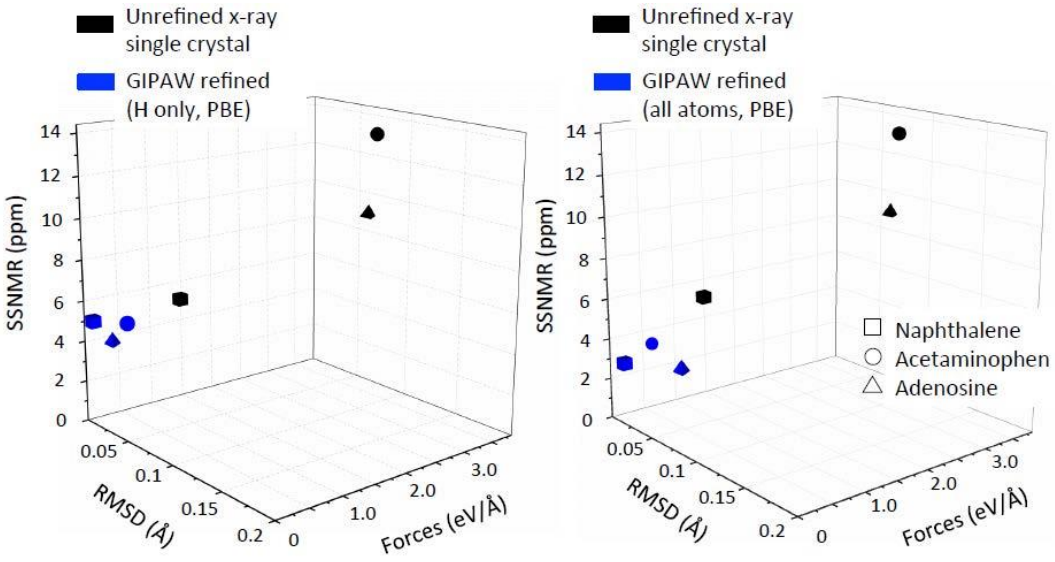


Figure 2-5: A GIPAW refinement of x-ray single crystal coordinates from model compounds was found to improve all figures of merit evaluated. The largest changes occurred upon refining hydrogen positions (left plot), however, refinement of all atoms (right plot) gave additional improvements showing that alteration of heavy atom positions is an important aspect of refinement.

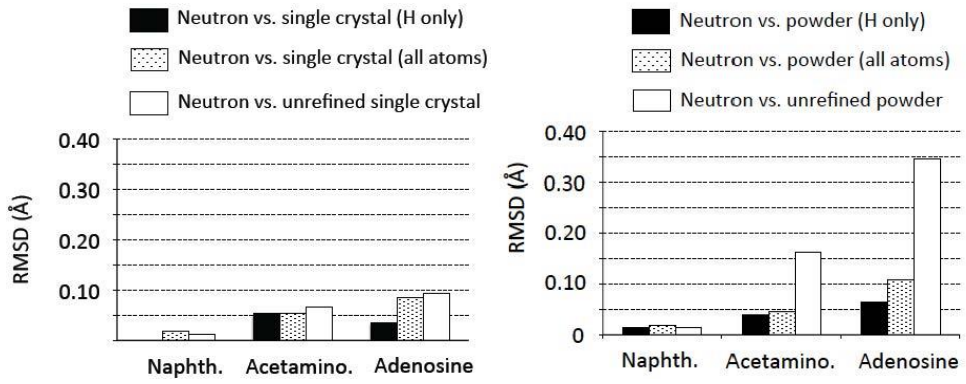


Figure 2-6: A comparison of RMSDs (Å) in atom positions showing neutron diffraction values versus both powder and single crystal x-ray coordinates after refinement of all atoms and after refinement of only H atoms. No hydrogen positions were reported for naphthalene's x-ray single crystal coordinates, thus these data are omitted.

Surprisingly, these refinement data show that the SSNMR fit for the single crystal x-ray structures before GIPAW refinement was worse than corresponding fit in powder diffraction structures in all cases. Most of the differences between the single crystal and

powder structures were found to lie in the hydrogen positions as demonstrated by refining all atoms and comparing to a separate refinement involving only hydrogens. In most cases, the largest improvements came from the hydrogen refinement. Presumably, hydrogen positions in powder diffraction structures are assigned based on external information (e.g. comparable neutron diffraction single crystal positions) and thus more accurate coordinates are reported. The GIPAW refinement of all x-ray single crystal structures produced coordinates that were statistically indistinguishable from similarly refined powder structures. Based on the fit to SSNMR data, all refined single crystal x-ray structures are significantly better than the initial diffraction structures at a statistical confidence of > 98%.

Changes in bond lengths from GIPAW refinement

As a final measure of the changes that occur upon GIPAW refinement, adjustments to bond lengths in the model compounds were examined. Bond length data are of particular interest because the majority of the SSNMR structural studies involve measurement of dipolar couplings and often report bond lengths derived from these data. For powder diffraction structures, the GIPAW refinement changed the C–C, C–N and C–O bond lengths, respectively, by ± 0.016 Å, ± 0.013 Å and ± 0.012 Å on average. Individual bond length changes ranged from ± 0.000 – 0.040 Å for these bonds. Bond lengths involving hydrogen atoms showed larger changes ranging from ± 0.030 – 0.101 Å with an average bond length change of ± 0.087 Å observed.

Refinement of the x-ray single crystal data caused smaller adjustments in heavy atom bond lengths with the C–C, C–N and C–O lengths changing, respectively, by ± 0.014 Å,

± 0.009 Å, and ± 0.004 Å on average. Changes in individual bond lengths varied from ± 0.000 – 0.026 Å. Bonds involving hydrogen exhibited significantly larger adjustments than corresponding bonds in powder structures, reflecting the differences in the methods used to locate hydrogens, as discussed previously. The average bond to a hydrogen lengthened by 0.138 Å upon refinement while changes to individual bonds varied from 0.025 - 0.218 Å. A comparison of these bond lengths to corresponding values from neutron diffraction lengths is illustrated in Figure 2-7. A more extensive comparison for each individual compound is included as Electronic Supplementary Information. These bond length data can be compared to SSNMR dipole coupling data where uncertainties typically range from ± 0.1 – 0.7 Å for X/Y sites (X & Y = ^{15}N, ^{13}C and ^{31}P)[49] and from ± 0.04 – 0.07 Å for ^{1}H/^{13}C pairs.[50] Thus, the changes in non-hydrogen positions from GIPAW refinement are roughly an order of magnitude smaller than the uncertainty in typical dipole coupling SSNMR measurements and would therefore not be detectable by such methods. However, many of the changes involving hydrogens would be observable by dipole coupling methods.

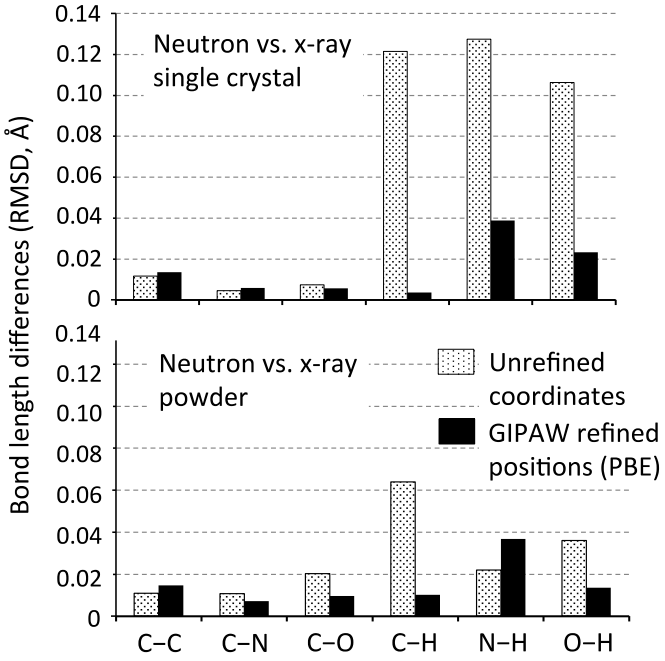


Figure 2-7: A comparison of differences (RMSD, Å) in bond lengths for structures obtained from x-ray single crystal and powder diffraction data versus single crystal neutron values. The influence of GIPAW refinement for each bond type is indicated on the horizontal axis. The number of C–C, C–N, C–O bonds included in this analysis was, respectively, 19, 12 and 7. The number of C–H, N–H, and O–H bonds evaluated was, respectively, 15, 3 and 4

Refinement of single crystal neutron diffraction coordinates

One of the most striking features of the bond length data in Figure 2-7 is that GIPAW adjustments to certain heavy atom bond lengths of either powder or single crystal x-ray diffraction data create bonds that are *less consistent* with the original neutron diffraction bond lengths. This is most frequently observed in C– C bond lengths where the refinement often lengthens the bonds beyond the corresponding neutron values. These changes can create novel structures that diverge from the original neutron diffraction single crystal coordinates. The crucial question is whether these new structures represent an improvement or a negligible modification that lies within the error of the original

coordinates. In all cases, the GIPAW refinements appear to genuinely improve structure based on the fact that statistically significant improvements are observed in both the forces upon the atoms and the agreement with experimental SSNMR shift tensor data. A comparison of these parameters is given in Figure 2-8 where refined and unrefined neutron structures are compared. In the cases of naphthalene and acetaminophen, the new coordinates are statistically distinct from the initial neutron structures based on improved agreement in the SSNMR fit. The refined adenosine coordinates, however, are statistically indistinguishable from the initial neutron structure based on a similar comparison of SSNMR shifts.

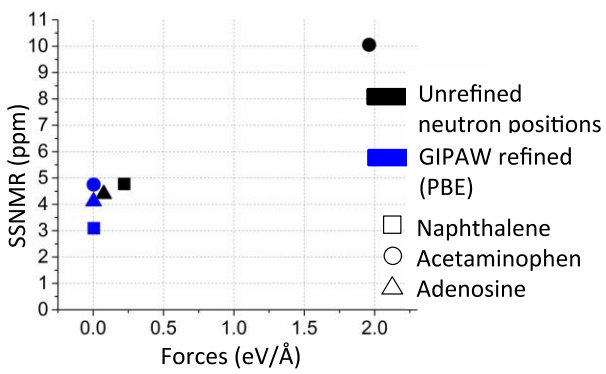


Figure 2-8: Comparison of forces and SSNMR fit in the unrefined single crystal neutron and the GIPAW refined neutron structures.

Refinement of single crystal neutron data with GIPAW has previously been demonstrated using ^{13}C tensor data consisting of six shift tensor values per nucleus.[20] The data collected in this study establishes that the three tensor principal values are sufficiently sensitive to also monitor improvements to neutron diffraction coordinates. It is notable that the changes in the neutron coordinates are extremely small and are comparable to the

error in coordinates found in many diffraction studies (Figure 2-9). For purposes of comparison, Figure 2-9 also includes differences between the neutron diffraction coordinates and the x-ray single crystal coordinates before refinement of either structure. Corresponding changes in non-hydrogen bond lengths upon refinement are very similar to those observed from GIPAW refinement of single crystal x-ray diffraction data. Specifically, the C–C, C–N, and C–O bonds in the neutron data changed, respectively, by ± 0.013 Å, ± 0.006 Å, and ± 0.005 Å on average. However, changes to bond lengths involving hydrogen were 2.6 times smaller than those observed in the single crystal x-ray data with an average change of only ± 0.053 Å observed. The changes in non-hydrogen bond lengths are most likely too small to be detected by dipole coupling methods. Adjustments involving hydrogens, however, are comparable to the error in dipole coupling data (i.e. ± 0.04 – 0.07 Å). Thus changes in neutron diffraction hydrogen positions may be detectable in favorable cases. In general, SSNMR tensor data appears to be uniquely capable of monitoring adjustments to heavy atom positions in single crystal neutron diffraction data.

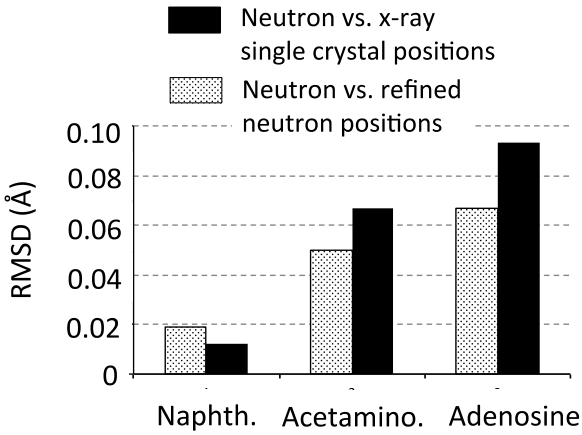


Figure 2-9: The RMSDs (Å) in atomic positions comparing the unrefined single crystal neutron coordinates with GIPAW refined neutron and unrefined x-ray single crystal positions. These RMSD values compare all atoms in the structures

Evaluating convergence in the GIPAW refined structures

The present study compares GIPAW refinement of crystal structures obtained from three different diffraction methods. One important consideration is whether the refined structures converge to a single structure for a given molecule. Three criteria were considered here consisting of forces on the atoms, RMS deviation in atomic positions and agreement between experimental and calculated SSNMR tensor data. These metrics are compared in Table 2-4 and demonstrate that all structures converged to a common structure regardless of the source of the original diffraction data. However, all the parameters in Table 2-4 are averages that include all positions in the model structures. In contrast, the SSNMR data allow the agreement at individual atomic positions to be considered. When such SSNMR tensor data are evaluated, the acetaminophen refined powder structure is found to be inferior to the other structures due to a poor fit at a single carbon (i.e. C8). This result demonstrates that the GIPAW refinements can sometimes fail to find the global minimum. Figure 2-10 shows that the three refined acetaminophen structures differ primarily at the methyl group, where differences in hydrogen atom positions are observed. Similar differences in converged structures have previously been noted by others[18c,19] and appear to be a fairly common outcome of this type of refinement. A complete list of computed ^{13}C tensor values for all converged structures is given as Electronic Supplementary Information.

Table 2-4: A comparison of several parameters to assess convergence in the GIPAW refined model compounds.

Structure	Source of data	SSNMR (ppm)[a]	Forces (eV Å^{-1})	RMSD (Å)[b]
Naphthalene	Powder (x-ray)	3.01	0.002	0.001
	X-ray single crystal	3.11	0.004	0.002
	Neutron single crystal	3.09	0.005	–
Acetaminophen	Powder (x-ray)	4.78	0.003	0.033
	X-ray single crystal	4.73	0.002	0.006
	Neutron single crystal	4.75	0.004	–
Adenosine	Powder (x-ray)	4.16	0.003	0.059
	X-ray single crystal	4.13	0.003	0.031
	Neutron single crystal	4.11	0.003	–

[a]These values are obtained by comparing principal values computed using the GIPAW refined coordinates to experimental principal value data using Alderman's icosahedral representation.[34] The PBE functional was used for all calculations.
[b]A comparison of all atoms in the GIPAW refined structures to corresponding sites in the GIPAW refined neutron single crystal structures.

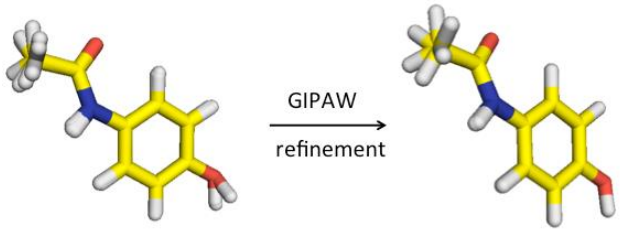


Figure 2-10: A comparison of geometry differences found in the three diffraction structures of acetaminophen showing the three superimposed crystal structures before (left) and after GIPAW refinement. In both cases the three molecules have been superimposed to minimize differences in heavy atom positions.

Overall, the results presented here suggest that crystal structure data from any source can be refined to a final structure with a quality rivaling data obtained from single crystal neutron diffraction. The present study is limited to three simple structures and extension of these methods to a greater variety of diffraction structures is needed to establish the general values of these refinements.

Experimental

The FIREMAT ^{13}C shift tensor data for acetaminophen were acquired on a CMX-400 Chemagnetics spectrometer operating at a 100.61916 MHz. The spectrum was acquired with TPPM ^{1}H decoupling at a frequency of 400.11880 MHz with a 9.1 μs 180° pulse and a phase modulation of ± 12° between consecutive pulses. Spectral widths of 26.5 kHz and 50.5 kHz were used in the evolution and acquisition dimensions, respectively. A total of 11 evolution increments of 192 scans each were collected using a 180 s recycle time for a total analysis time of 4.4 days. Data processing was completed according to a process described elsewhere.[34] A sample spinning speed of 2405 Hz was used and the spectrum was externally referenced to the methyl resonance in hexamethyl benzene at 17.35 ppm.

The $^{1}H/^{13}C$ heteronuclear correlation spectrum of acetaminophen was acquired on a Bruker 600 MHz wide bore spectrometer operating at a ^{13}C frequency of 150.919 MHz using a low-E 3.2 mm magic-angle spinning probe[51] and a spinning frequency of 10 kHz. A cross-polarization time of 50 μs was employed for the $^{1}H \rightarrow ^{13}C$ HETCOR transfer. A total of 100 (50 Complex) evolution increments of 4 scans each were acquired with a 180 s recycle delay. Two-step super-cycled phase modulated Lee-Goldburg homonulear decoupling[52] was applied during ^{1}H evolution. A glycine sample was used for the optimization and calibration of ^{1}H chemical shift scaling and referencing. Digital resolutions of 24.4 and 158.0 Hz per point were acquired in the acquisition and evolution dimensions, respectively.

The total experiment time was 20 h. The acetaminophen sample used for all measurements was purchased from Aldrich and used without further purification.

Three polymorphs of acetaminophen are known.[53] The phase used in all SSNMR analyses was found to be the $P2_1/c$ form based on x-ray powder diffraction analysis of the acetaminophen powder. Data were collected at room temperature using an X'Pert PRO diffractometer (PANalytical) with Cu K alpha radiation (λ = 1.5406 Å) and compared to a the experimental powder patterns[53] to establish the phase. Data were collected from 2.0–70.0° using a scan step size of 0.013°. A total of 5229 points were collected. Diffraction data from the $P2_1/c$ polymorphs were used in all GIPAW calculations reported here. Only one phase is known for adenosine ($P2_1$) and naphthalene ($P2_1/c$) and thus independent verification of the solid form studied here was not considered necessary.

All calculations were completed on the Lionxj metacluster of the Research Computing and Cyperinfrastructure at Pennsylvania State University. The GIPAW calculations were performed with software from Accelrys Software Inc.[54] using the CASTEP module. The periodic structure of the lattice is intrinsically incorporated in all GIPAW calculations. Unit cell dimensions were not refined in any of the computations because these parameters are systematically overestimated by the methods used.[18a] The PBE,[40] PW91,[41] and LDA-CA-PZ[42] functionals were employed along with ultrasoft pseudopotentials. The so-called "ultra-fine" level was selected for the CASTEP parameters. The ultra-fine option uses planewave basis set cut-off energies of 390 eV for hydrocarbons and 550 eV if oxygen is present. The threshold convergence for SCF tolerance was 5×10^{-7} eV/atom. A k-point spacing of 0.071/Å was employed. The minimizing approach of

Broyden, Fletcher, Goldfarb and Shanno[55] was used for geometry optimizations. For convergence of the optimization, a change in energy threshold of 5×10^{-6} eV/atom was reached, a maximum Cartesian force of 0.01 eV/Å on all atoms was achieved, and the maximum displacement of 5×10^{-4} Å for each atom was observed.

A conversion of computed shielding to shift is necessary before a comparison can be made to experimental values. In the most extensive evaluation of ^{13}C tensor data to date, over 600 tensor values from 14 compounds were computed and compared with experimental data.[20] In this study a process was recommended that involved fitting a least squares straight line to a plot of shielding versus shift and using the fit slope and intercepts to convert shielding values to shifts. The alternative of using a slope equal to one and an intercept that corresponds to a reference compound were shown to introduce large systematic errors. An extensive discussion of the issues involved in the conversion of shielding values is summarized elsewhere and is beyond the scope of this study.[20] To ensure consistency with the previous ^{13}C work conversion factors for the GGA-PBE, the GGA-PW91 and LDA-CA-PZ functionals were determined by computing shielding values for the 14 organic crystals previously studied. Each of these structures has a known neutron diffraction structure and previously reported ^{13}C shift tensor values.[20] This conversion differs from that previously reported in that the three principal shift values per nucleus were employed here rather than the full tensor (six shift per nucleus) used previously. The compounds selected consist of a variety of carbohydrates and aromatic hydrocarbons. Shielding values were computed using both the reported neutron diffraction coordinates and the same coordinates after GIPAW refinement (GGA-PBE). In both cases the PBE

functional was found to provide the best fit to experimental SSNMR data. The GIPAW structural refinement was found to decrease the error in computed tensor values in all 14 compounds, an improvement that is statistically significant at the 95% confidence level. The slopes, intercepts, and errors for all computed values after the GIPAW refinement of the neutron diffraction coordinates are summarized in Table 2-5. The conversion between GGA-PBE shieldings and shift for the refined neutron coordinates was found to be: shift = (shielding – 171.41)/(-1.0316). This conversion equation was used for all calculations reported herein.

Table 2-5: Slope, intercept and RMS deviation between computed and experimental ^{13}C shift tensors for different theoretical methods.

Treatment	Error	Slope	Intercept
GGA-PBE/Ultrafine	3.31	1.0316	171.41
GGA-PW91/Ultrafine	3.43	1.0285	170.99
LDA-CA-PZ/Ultrafine	3.58	1.0588	171.90

Conclusion

Many of the key results from this study are summarized in Figures 2-4, 2-5, 2-8 and 2-9. In these figures it is notable that the RMSD values change by only minor amounts upon refinement. In fact, in most cases, the RMSD changes are comparable to the experimental error in atom positions, making the structural modifications statistically indistinguishable from the unrefined coordinates. This insensitivity is particularly problematic considering that RMSD is the figure-of-merit most commonly reported when comparing two structures. In contrast, the forces on the individual atoms are highly sensitive to GIPAW refinement

with changes of 2–3 orders of magnitude commonly observed. Force has the limitation that corresponding experimental data are not available for comparison. This can be problematic when several crystal structures of the same compound are being considered since refinement often creates more than one final conformation. In this case, distinguishing between final structures by forces is not possible since all have acceptable values. This structural ambiguity was observed here in acetaminophen and has been previously noted by others.[18c,19] The final parameter considered, SSNMR ^{13}C tensor principal values, is also quite sensitive to refinement and has the advantage that corresponding experimental data are available to verify results. The SSNMR data are unique among the parameters studied in that data is available for each individual atomic site in a molecule. This "site-resolved" feature may prove useful in resolving the GIPAW structural ambiguities mentioned above for forces. For example, by providing experimental data for comparison at conformationally ambiguous atomic sites SSNMR may allow one model to be selected from among several candidates created by conventional GIPAW refinements. Thus force and SSNMR data can be used in a complimentary manner. Currently, the refinements described herein do not provide information on distortions due to thermal motions or other movements in the lattice. Inclusion of such effects is now possible[56] and may prove to be an important consideration in future work.

Application of GIPAW refinements to a wide variety of crystal structures is now being pursued in our laboratory. These analyses indicate that improvements in the SSNMR ^{13}C tensor fit from refinement are usually significantly larger than those reported in the present study. This new work includes x-ray powder and single crystal structures and

indicates that in many diffraction structures, SSNMR fit improves by a factor of 3–5 upon refinement. Future work will therefore emphasize refinement of a more diverse collection of structures including structures from lower resolution techniques where larger improvements can be expected. For example, electron diffraction structures are frequently of such low resolution that some atomic positions cannot be determined. Likewise, the accuracy of crystal structures obtained solely from SSNMR cannot be evaluated by the same metrics used for diffractions structures, making comparison with related diffraction structures difficult. This study suggests that further refinement of such data will result in coordinates comparable to or surpassing the quality of data obtained from single crystal neutron diffraction.

Acknowledgements

This work was supported, in part, by Pennsylvania State University's Materials Simulation Center (Materials Research Institute), Research Computing and Cyberinfrastructure (Information Technology Services, http://rcc.its.psu.edu) and the Center for Nanoscale Science. Drs. Evelyn Moreno-Calvo and Miquel Angel Cuevas-Diarte (Universidad de Barcelona) acquired the X-ray powder diffraction data for acetaminophen that was used to verify the polymorphic form of the material studied herein. A portion of this work was performed at the National High Magnetic Field Laboratory, which is supported by National Science Foundation Cooperative Agreement No. DMR-1157490, the State of Florida, and the U.S. Department of Energy. We are grateful to Drs. Zhehong Gan and Ivan Hung (National High Magnetic Field Laboratory) for acquisition of the $^{1}H/^{13}C$ correlation data for acetaminophen.

References

1. W. L. Bragg, *Proc. R. Soc. London Ser. A*, 1913, **88**, 428-438.

2. R. Brill, H. G. Grimm, C. Hermann and C. Peters, *Ann. Phys.*, 1939, **34**, 393-445.

3. R. D. Shannon and C. T. Prewitt, *Acta Crystallogr. Sect. B,* 1969, **25**, 925-946.

4. J. M. Bijvoet, A. F. Peerdeman and A. F. van Bommel, *Nature*, 1951, **168**, 271-272.

5. J. S. Miller and J. J. Novoa, *Acc. Chem. Res.* 2007, **40**, 189-196.

6. R. Taylor and O. Kennard, *J. Am. Chem. Soc.*, 1982, **104**, 5063-5070.

7. W. Steurer and S. Deloudi, *Acta Crystallogr. Sect. A*, 2008, **64**, 1-11.

8. (a) F. Taulelle and C. Huguenard, *Stud. Surf. Sci. Catal.*, 2001, **135**, 1414-1421. (b) J. Dutour, N. Guillou, C. Huguenard, F. Taulelle, C. Mellot-Draznieks and G. Ferey, *Solid State Sci.*, 2004, **6**, 1059-1067. (c) I. J. King, F. Fayon, D. Massiot, R. K. Harris and J. S. O. Evans, *Chem. Commun.*, 2001, **18**, 1766-1767.

9. (a) D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti and I. K. Zhitkov, *Acta Crystallogr. Sect. B*, 2011, **67**, 535-551. (b) G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B.

Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Crystallogr. Sect. B*, 2009, **65**, 107-125. (c) G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt and P. Verwer, *Acta Crystallogr. Sect. B*, 2005, **61**, 511-527. (d) W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallogr., Sect B* **2002**, *58*, 647-661. (e) J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammin, J. D. Dunitz, A. Gavezzotti, D. W. M. Hoffmann, F. J. J. Leusen, W. T. M. Mooji, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallogr., Sect. B*, 2000, **56**, 697-714.

[10]. (a) D. H. Brouwer, In *NMR Crystallography*, ed. R. K. Harris, R. E. Wasylishen, and M. J. Duer, Wiley, Chichester, 2009, p. 263-274. (b) K. D. M. Harris, X. Mingcan, In *NMR Crystallography*, ed. R. K. Harris, R. E. Wasylishen, and M. J. Duer, Wiley, Chichester, 2009, p. 275-287.

[11]. (a) C. A. Fyfe, D. H. Brouwer, A. R. Lewis, L. A. Villaescusa and R. E. Morris, *J. Am. Chem. Soc.*, 2002, **124**, 7770-7778. (b) C. A. Fyfe, D. H. Brouwer, A. R. Lewis and J. –

M.Chézeau, *J. Am. Chem. Soc.*, 2001, **123**, 6882-6891. (c) D. H. Brouwer, I. L. Moudrakovski, K. A. Udachin, G. D. Enright and J. A. Ripmeester, *Cryst. Growth Des.*, 2008, **8**, 1878-1885.

[12]. S. Meejoo, B. M. Kariuki, S. J. Kitchin, E. Y. Cheung, D. Albesa-Jové and K. D. M. Harris, *Helv. Chim. Acta*, 2003, **86**, 1467-1477.

[13]. M. Tremayne, B. M. Kariuki and K. D. M. Harris, *Angew. Chem. Int. Ed. Engl.*, 1997, **36**, 770-772.

[14]. M. Edgar, V. J. Carter, D. P. Tunstall, P. Grewal, V. Favre-Nicolin, P. A. Cox, P. Lightfoot and P. A. Wright, *Chem. Commun.*, 2002, **8** 808-809.

[15]. (a) R. Witter, U. Sternburg, S. Hesse, T. Kondo, F. T. Koch and A. S. Ulrich, *Macromolecules*, 2006, **39**, 6125-6132. (b) D. H. Brouwer, *J. Am. Chem. Soc.*, 2008, **130**, 6306-6307. (c) J. K. Harper, D. M. Grant, Y. Zhang, P. L. Lee and R. B. Von Dreele, *J. Am. Chem. Soc.*, 2006, **128**, 1547-1552. (d) J. K. Harper, J. A. Doebbler, E. Jacques, D. M. Grant and R. B. Von Dreele, *J. Am. Chem. Soc.*, 2010, **132**, 2928-2937. (e) D. H. Brouwer, D. H. *J. Magn. Reson.*, 2009, **194**, 136-146.

[16]. C. J. Pickard and F. Mauri, F. *Phys. Rev., B*, 2001, **63**, 245101.

[17]. J. Hutter, P. Ballone, M. Bernasconi, P. Focher, E. Fois, S. Goedecker, M. Parrinello, M. Tuckerman, CPMD, 3.7 ed.; MPI für Festkörperforschung and IBM Zurich Research Laboratory, 1997-2001.

[18]. (a) S. E. Ashbrook, M. Cutajar, C. J. Pickard, R. I. Walton and S. Wimperis, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5754-5764. (b) M. Profeta, F. Mauri and C. J. Pickard, *J. Am.*

Chem. Soc., 2003, **125**, 541-548. (c) D. H. Brouwer, I. L. Moudrakovski, R. J. Darton and R. E. Morris, *Magn. Reson. Chem.*, 2010, **48**, S113-S121.

[19]. E. Salager, R. S. Stein, C. J. Pickard, B. Elena and L. Emsley, *Phys. Chem. Chem. Phys.*, 2009, **11**, 2610-2621.

[20]. J. C. Johnston, R. J. Iuliucci, J. C. Facelli, G. Fitzgerald and K. T. Mueller, *J. Chem. Phys.*, 2009, **131**, 144503-144514.

[21]. (a) D. H. Brouwer, *J. Magn. Reson.*, 2008, **194**, 136-146. (b) E. Moreno-Calvo, G. Gbabode, R. Cordobilla, T. Calvert, M. A. Cuevas-Diarte, P. Negrier and D. Mondieig, *Chem. Euro. J.*, 2009, **15**, 13141-13149.

[22]. F. Taulelle, In *NMR Crystallography*, ed. R. K. Harris, R. E. Wasylishen, and M. J. Duer, Wiley, Chichester, 2009, p. 245-262.

[23]. R. A. Olsen, J. Struppe, D. W. Elliott, R. S. Thomas, L. J. Mueller, *J. Am. Chem. Soc.* 2003, **125**, 11784-11785.

[24]. C. Luchinat, G. Parigi, E. Ravera, M. Rinaldelli, *J. Am. Chem. Soc.* 2012, **134**, 5006-5009.

[25]. C. W. Padgett, H. D. Arman, and W. T. Pennington, *Cryst. Growth Des.*, 2007, **7**, 367-372.

[26]. K. H. Stone, S. Lapidus, P. W. Stephens, *J. Appl. Crystallogr.*, 2009, **42**, 385-391.

[27]. R. Caminiti, G. Ortaggi, R. A. Mazzei, P. Ballirano and R. Rizzi, *Powder Diffr.*, 2000, **15**, 108-111.

[28]. D. W. J. Cruickshank, *Acta Crystallogr.*, 1957, **10**, 504-508.

[29]. M. Haisa, S. Kashino, R. Kawai and H. Maeda, *Acta Crystallogr., Sect. B*, 1976, **32**, 1283-1285.

[30]. T. F. Lai and R. E. Marsh, *Acta Crystallogr., Sect. B*, 1972, **28**, 1982-1989.

[31]. S. C. Capelli, A. Albinati, S. A. Mason and B. T. M. Willis, *J. Chem. Phys. A*, 2006, **110**, 11695-11703.

[32]. C. C. Wilson, *Z. Kristallogr.*, 2000, **215**, 693-701.

[33]. W. T. Klooster, J. R. Ruble, B. M. Craven and R. K. McMullan, *Acta Crystallogr., Sect. B*, 1991, **47**, 376-383.

[34]. D. Stueber and D. M. Grant, *J. Am. Chem. Soc.*, 2002, **124**, 10539-10551.

[35]. D. W. Alderman, G. McGeorge, J. Z. Hu, R. J. Pugmire and D. M. Grant, *Mol. Phys.*, 1998, **95**, 1113-1126.

[36]. M. S. Sherwood, J. C. Facelli, D. W. Alderman and D. M. Grant, *J. Am. Chem. Soc.*, 1991, **113**, 750-753.

[37]. J. C. Facelli, D. M. Grant, *Nature*, 1993, **365**, 325–327.

[38]. N. R. Jagannathan, *Current Sci.*, 1987, **56**, 827-830.

[39]. S. J. Cark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, M. C, Payne, *Z. Kristallogr. - New Cryst. Struct.* 2005, **220**, 567-570.

[40]. J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865-3868.

[41]. J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1992, **46**, 6671-6687.

[42]. D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.*, 1980, **45**, 566-569.

[43]. Representative examples include: (a) J. Trotter, *Acta Crystallogr.*, 1960, **13**, 732-736. (b) L. J. Fitzgerald, R. E. Gerkin, *Acta Crystallogr., Sect. C*, 1993, **49**, 1952-1958. (c) F. R. Fronczek, G. Gannoch, W. L. Mattice, F. L. Tobiason, J. L. Broeker and R. W.

Hemingway, *J. Chem. Soc., Perkin Trans. 2*, 1984, 1611-1616. (d) A. L. Spek, B. Kojic-Prodic, and R. P. Labadie, *Acta Crystallogr., Sect. B*, 1984, **40**, 2068-2071.

[44]. S. C. Abrahams, J. M. Robertson and J. G. White, *Acta Crystallogr.*, 1949, **2**, 233-238.

[45]. (a) F. R. Ahmed and D. W. J. Cruickshank, *Acta Crystallogr.*, 1952, **5**, 852-853. (b) V. I. Ponomarev, O. S. Filipenko and L. O. Atoumyan, *Kristallografiya (Crystallogr. Rep.)*, 1976, **21**, 392-395. (c) N. B. Chanh and Y. Haget-Bouillaud, *Acta Crystallogr., Sect. B*, 1972, **28**, 3400-3404. (d) C. P. Brock and J. D. Dunitz, *Acta Crystallogr. Sect. B*, 1982, **38**, 2218-2228. (e) H. C. Alt and J. Kalus, *Acta Crystallogr., Sect. B*, 1982, **38**, 2595-2600. (f) J. Oddershede and S. J. Larson, *J. Phys. Chem. A*, 2004, **108**, 1057-1063. (g) F. P. A. Fabbiani, D. R. Allan, S. Parsons and C. R. Pulham, *Acta Crystallogr., Sect. B*, 2006, **62**, 826-842.

[46]. (a) G. S. Pawley and E. A. Yeats, *Acta Crystallogr., Sect. B*, 1969, **25**, 2009-2013. (b) I. Natkaniec, A. V. Belushkin, W. Dyck, H. Fuess and C. M. E. Zeyen, *Z. Krystallogr.*, 1983, **163**, 285-293.

[47]. P. J. Byrne, J. E. Warren, R. E. Morris and S. E. Ashbrook, *Solid State Sci.*, 2009, **11**, 1001-1006.

[48]. (a) M. Strohmeier, D. W. Alderman and D. M. Grant, *J. Magn. Reson.*, 2002, **155**, 263-277. (b) D. H. Barich, J. S. Clawson, D. Stueber, M. Strohmeier, R. J. Pugmire and D. M. Grant, *J. Phys. Chem. A*, 2002, **106**, 11379-11379.

[49]. J. M. Griffiths and R. G. Griffin, *Anal. Chim. Acta*, 1993, **283**, 1081-1101.

[50]. B. –J. van Rossum, C. P. de Groot, V. Ladizhansky, S. Vega and H. J. M. de Groot, *J. Am. Chem. Soc.*, 2000, **122**, 3465-3472.

[51]. Gor'kov, P. L.; Chekmenev, E. Y.; Li, Cotton, M.; Buffy, J. J.; Traaseth, N. J.; Veglia, G.; Brey, W. W. *J. Magn. Reson.* **2007**, *185*, 77-93.

[52]. Vinogradov, E.; Madu, P. K.; Vega, S. *Chem. Phys. Lett.* **2002**, *354*, 193-202.

[53]. (a) M. Haisa, S. Kashino and H. Maeda, *Acta Crystallogr. Sect. B: Struct. Crystallogr. Cryst. Chem.*, 1974, **30**, 2510-2512. (b) M. Haisa, S. Kashino, R. Kawai, and H. Maeda, *Acta Crystallogr. Sect. B: Struct. Crystallogr. Cryst. Chem.* 1976, **32**, 1283-1285. (c) M. –A. Perrin, M. A. Neumann, H. Elmaleh, and L. Zaske, *Chem. Commun.* 2009, 3181-3183.

[54]. Accelrys Software Inc., Materials Studio, Release 6.0, San Diego, 2011.

[55]. B. G. Pfrommer, M. Cote, S. G. Louie, and M. L. Cohen, *J. Comput. Phys.*, 1997, **131**, 233-240.

[56]. J. –N. Dumez and C. J. Pickard, *J. Chem. Phys.*, 2009, **130**, 104701.

CHAPTER 3: ESTABLISHING ACCURATE HIGH-RESOLUTION CRYSTAL STRUCTURES IN THE ABSENCE OF DIFFRACTION DATA AND SINGLE CRYSTALS - A SOLID STATE NMR APPROACH

Reproduced from Kalakewich, K. et al. Establishing Accurate High-Resolution Crystal Structures in the Absence of Diffraction data and Single Crystals – a Solid State NMR approach. *Cryst. Growth Des.* 2013, 13, 5391. Copyright 2013. With permission from the American Chemical Society

Keyton Kalakewich,[a] Robbie J. Iuliucci,[b] James K. Harper[a,*]

[a]*The University of Central Florida, Department of Chemistry, Orlando FL 32816 and* [b]*Washington and Jefferson College, Department of Chemistry, Washington, PA 15301.*

Abstract

Predicting accurate crystal structures from theoretical consideration has proven to be remarkably challenging. Although significant progress has been made and numerous approaches have now been investigated, selection of the correct structure as the first choice in blind studies is still rarely achieved. Here a process is described that consistently identifies the correct structure from the myriad candidates created from typical crystal structure prediction software. This approach relies on ^{13}C solid-state NMR data and a secondary refinement process that includes lattice fields. Four structures are considered and in all cases the correct structure is selected as the first choice and the only statistically feasible candidate. Data from ^{13}C chemical shift tensor principal values are found to provide better selectivity, but ^{13}C isotropic shifts also consistently identify the correct structure. This process involves only experimental NMR data and computer-generated structures, yet the structures created appear to rival the accuracy of structures derived from single crystal diffraction methods including single crystal neutron diffraction.

Introduction

The 1912 observation of x-ray diffraction[1] has proven to be one of the key experimental findings of the past century. This discovery made possible the first crystal structure determinations[2] and provided the first view of matter at the atomic level. Insights gained from crystallography provided key guiding principles in the development of modern chemistry and solid-state physics and served as a crucial tool in the development of molecular biology.[3] Despite the remarkable success of crystallography, many materials remain unsuitable for diffraction methods due to difficulty in growing single crystals for diffraction. The development of modern powder diffraction structure determination methods has overcome some of these challenges and structures can now be obtained for many materials that fail to form suitable crystals,[4] including some macromolecules.[5,6] However, powder methods are usually restricted to smaller molecules where overlap of Bragg peaks is limited. Powder methods can also have difficulty treating compounds that diffract poorly such as organic materials containing only light atoms. These challenges to conventional diffraction methods emphasize the need to develop alternative techniques.

Over the past decade, the inclusion of solid-state NMR (SSNMR) data in crystal structure determination has become a topic of considerable interest and has proven effective in characterizing certain materials that have been difficult or impossible to characterize by other techniques.[7] Such work is now widely referred to as NMR crystallography and is often defined as analysis of microcrystalline powders that combine information from diffraction studies (e.g. space group) with structural models from SSNMR data. In contrast to this approach, an alternative was proposed in 2006 that requires *no*

diffraction data, single crystals or knowledge of space group and instead uses experimental SSNMR data to rank a wide variety of candidate structures created from a computational process known as crystal structure prediction (CSP).[8] Modern CSP methods can create thousands of candidate structures solely from knowledge of atomic connectivity with some programs capable of predicting structures in all space groups.[9] Historically, identification of the correct structure has been very challenging. The 2006 study reduced the number of feasible crystal structures by nearly an order of magnitude over comparisons based on lattice energy. Candidates were ranked by agreement between experimental ^{13}C tensor shifts and corresponding theoretical values computed for each structure. Although this CSP-NMR study correctly identified the crystal structure in one case and ranked the correct structure among the top five in all cases, an ideal technique should consistently identify the correct structure. One possible limitation in this early study was that lattice fields were neglected in calculated NMR data.

Recently, several approaches have been introduced that allow lattice factors from periodic crystal structures to be included in calculations.[10,11] One of the more widely available algorithms is the "gauge including projector augmented wave" (GIPAW) method.[10] The aim of the present work is to extend the ability of ^{13}C NMR tensor data to select a correct structure from CSP candidates by including lattice factors. This process has recently been used with GIPAW and ^{1}H isotropic shifts and shown to correctly identify the crystal structure of thymol.[12] Tensor data are usually considered to be more sensitive to structure than isotropic shifts and are therefore expected to improve the analysis. In the

following, we show that inclusion of GIPAW together with a crucial structural refinement step allows the correct structure to always be selected as the best fit.

Experimental Section

All GIPAW calculations were performed using CASTEP,[13] which intrinsically includes lattice structure. The computations employed the PBE functional with ultrasoft pseudopotentials, a planewave basis set cut-off energy of 550 eV and were run using the "ultra-fine" level. Other parameters included a k-point spacing of 0.071 $Å^{-1}$ and a threshold convergence for SCF tolerance of 5×10^{-7} eV $atom^{-1}$. Geometry optimizations were performed by the minimizing method of Broyden, Fletcher, Goldfarb and Shanno.[14] Optimizations were considered to be converged when the energy change was less than 5×10^{-6} eV $atom^{-1}$, Cartesian forces on all atoms was less than 0.01 eV $Å^{-1}$, and the maximum displacement of each atom was less than 5×10^{-4} Å. Unit cell dimensions were not altered in refinements since PBE is known to systematically overestimate cell dimensions.[23]

Experimental ^{13}C tensor data were obtained from a prior study using large single crystals and a 2D chemical shift correlation experiment.[15] All data were acquired at ambient temperature and the error in individual tensor shifts was estimated to be $\pm$ 0.6 ppm. The procedure used to extract tensor data from the 2D chemical shift correlation data and to estimate errors are given elsewhere.[16] Although these data provided the complete tensor (i.e. six shifts per nucleus) for each carbon, only the three tensor principal values per nucleus were used in the present study to allow comparison with microcrystalline powders where principal values can be more routinely obtained. Crystal

structure prediction coordinates were generated using the program UPACK as described elsewhere.[21] Neutron diffraction structures are available for each of the four structures studies herein[17] and were used for structural comparisons. All neutron structures were obtained at room temperature.

A conversion of computed ^{13}C shielding to shift is required before a comparison can be made to experimental data. Here shielding was plotted versus shift and a least-squares fit of slope and intercept calculated to allow optimal conversion of shielding to shift. When the GIPAW refined neutron coordinates were used, all data points were found to belong to the same population and the conversion between shielding and shift was given by shift=(shielding–178.089)/(-1.1412). All comparisons were made in the icosahedral representation[18] using the shielding/shift conversion listed above. It is notable that polycyclic aromatics are known to have slopes and intercepts that differ from those derived for carbohydrates[19] and it is possible that other classes of compounds will have similar differences. Thus, the equation derived here may not be suitable for all materials and there is ambiguity regarding how to proceed when new structures are treated. An alternative suggested by Olejniczak et al. is to plot computed shielding versus experimental shifts for all model structures considered, then fit each with an adjustable slope and intercept.[20] The model having the largest coefficient of determination (i.e. R^2) is retained as the most probable structure.

Results and Discussion

The CSP data considered here were generated using a process described elsewhere[21] and include four simple carbohydrates: methyl α-D-glucopyranoside (**1**), methyl α-D-mannopyranoside (**2**), methyl α-D-galactopyranoside (**3**) and methyl β-D-xylopyranoside (**4**). This particular group was selected over a more structurally diverse set of compounds because unusually accurate ^{13}C tensor data are available for each structure from single crystal NMR studies.[15] Likewise, high quality neutron diffraction structures are known for each structure, allowing a comparison of the accuracy of the CSP structures selected. In the original study, 422 CSP structures were retained as feasible candidates based on calculated lattice energy. Of these, 367 (87.0%) were eliminated based on poor fit to experimental NMR ^{13}C tensors. The 55 remaining structures included 30 structures of **1**, 7 of **2**, 8 of **3** and 10 of **4**. All GIPAW calculations described herein were performed using only these 55 particularly challenging structures. Admittedly, reanalysis of the entire set of 422 structures with lattice including methods would likely change the set of structures retained. However, this is not a particularly important consideration since the set of retained structures will typically vary between different CSP software packages. The crucial consideration is whether the correct structure is included and if it can be distinguished from the other retained structures. Since the four correct structures were among the 55 candidates retained here, we consider this subset of data to be an adequate test of the ability to select a correct structure using NMR data and lattice including methods. For each structure, ^{13}C shift tensors were computed to assess the contribution of lattice factors in calculated shifts. All data were then compared to experimental data.

Rankings of the correct structure are summarized in 3-1 and compared with previously reported rankings that omit lattice.[8]

Table 3-1: Ranking of the NMR fit of the CSP structures matching the known neutron diffraction structure.

Structure	GIPAW[a]	Lattice omitted
1	1st of 30	1st of 30
2	7th of 7	3rd of 7
3[b]	5th of 8	5th of 8
4	7th of 10	5th of 10

[a]GIPAW calculations include lattice fields, but retain original CSP coordinates.
[b]The original CSP structure had an error in the C3OH hydrogen position, which was corrected before ranking the structures.

It is clear from Table 3-1 that the inclusion of lattice factors alone does not improve rankings in any of the structures. Yet it is known that inclusion of lattice does, in fact, improve agreement between computed and experimental ^{13}C tensors in related carbohydrates by nearly 1 ppm when accurate coordinates are available.[19] Thus, it is likely that structural errors are present in the CSP structures and that further refinement is possible. Structural refinement under lattice constraints has previously demonstrated the ability to improve structures from single crystal[19,22] and powder diffraction.[23] These refinements improve both hydrogen and non-hydrogen positions. Accordingly, the 55 CSP structures were further refined within the lattice fields. Following refinement, ^{13}C shift tensors were computed using GIPAW and compared to experimental data. Remarkably, when refinement was included, *the correct structure was always ranked as the best fit to the NMR data*. Figure 3-1 shows the NMR agreement before and after refinement for **4**.

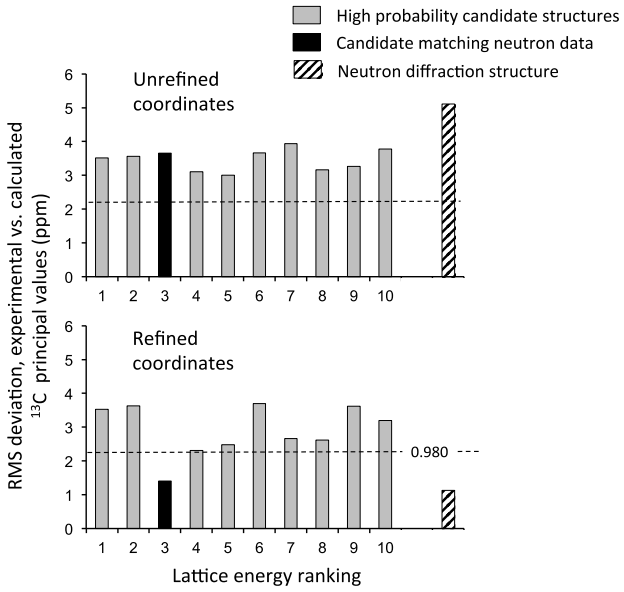


Figure 3-1: Agreement between experimental and calculated ^{13}C NMR shift tensor principal values for the 10 CSP structures of **4** that were retained in prior analysis.[8] Simply including lattice fields when calculating ^{13}C values (top) but not altering coordinates gave *no improvement* in the ability to select a correct structure. However, adjusting atomic positions by a refinement that includes lattice fields corrected minor bond length and conformational errors and allowed the correct structure to be selected unambiguously (bottom). All other structures were rejected with greater than 98% confidence. Notably, the NMR agreement for the neutron diffraction structure also improved significantly and differs from the initial neutron structure with > 99.9% confidence.

Similar lattice including refinement of the CSP structures of **1**, **2**, and **3** also gave improvements in selectivity relative to the unrefined coordinates. In each case the correct CSP structure was ranked as the best NMR fit after refinement (Figure 3-2) and the only statistically feasible candidate. In contrast, lattice energy ranks the correct structure as the best fit in only half of the molecules. Thus, NMR data offers a significant improvement in the ability to select the correct crystal structure from CSP.

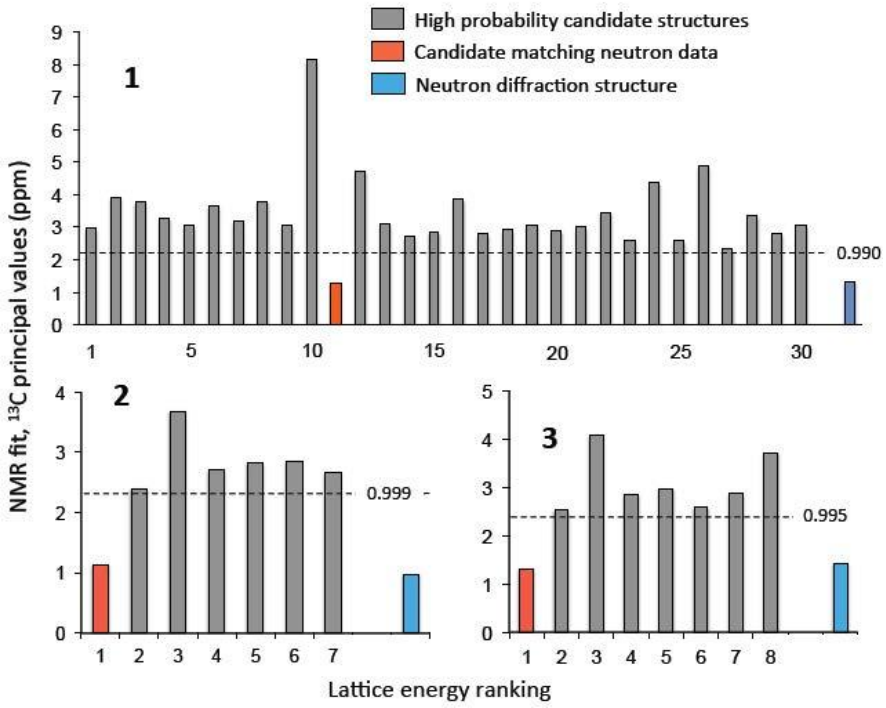


Figure 3-2: Agreement between experimental and computed NMR tensors for structures **1**, **2**, and **3** after a lattice-including refinement of all coordinates. Lattice energy rankings are shown on the horizontal scale. In each case the candidate crystal structure matching neutron diffraction data gave the best agreement with NMR ^{13}C shift tensors with all other structures rejected with ≥ 99.0% confidence. Statistically significant improvements were observed in neutron diffraction coordinates in all cases.

These data allow us to inquire whether inclusion of the lattice fields is more important in the structural refinement or in the calculation of the NMR parameters. Here, improvements of roughly 2–4 ppm were observed in calculated principal values when lattice fields were included in both the optimization and NMR calculations. By comparison, inclusion of only lattice factors in NMR calculations with no geometry optimization has been found to improve ^{13}C principal values by about 1 ppm in similar carbohydrates.[19] Thus the majority of the improvement appears to come from the geometry refinement rather than the inclusion of lattice in computed NMR parameters. Other nuclei may depend

more strongly on lattice in computed NMR shifts. However, even in these cases the improvements from geometry refinement may be important in distinguishing between candidate structures.

Recent work combining CSP, SSNMR data and GIPAW calculations has demonstrated that ^{1}H isotropic shift data also allow the correct crystal structure to be identified in thymol.[12] This innovative approach has the remarkable feature of being able to select the correct structure even when lattice energy fails to do so. However, it also has the limitation that several incorrect candidate structures cannot be eliminated with the high statistical confidence observed with ^{13}C principal values. The ^{1}H study also reported that ^{13}C isotropic shifts were unable to select the correct structure with the same confidence as ^{1}H shifts. This study did not include a lattice including refinement step. Given the influence such refinements have on structural selection here, it was deemed worthwhile to reinvestigate the selectivity of ^{13}C isotropic shifts using structures **1**–**4** after refinement.

Figure 3-3 shows the agreement between experimental and computed ^{13}C isotropic shifts for all refined CSP structures. In each case the correct crystal structure gave the best agreement with experimental data, illustrating that lattice including refinement is an important factor in identifying the correct crystal structure. It is notable that the selection using ^{13}C isotropic shift suffers from the same limitation observed in the ^{1}H study, namely that incorrect candidates are occasionally retained. Admittedly, the three incorrect structures retained in Figure 3-3 *can* be rejected at confidence levels of ≥ 88%, yet for *de novo* structural predictions, a very high confidence in the correct structure is required and this ambiguity could become problematic. Here all structures within 2 standard deviations

of the global error were retained (i.e. 95.4% confidence that rejected values are, in fact, incorrect). In contrast to this limitation to isotropic shifts, ^{13}C principal values do provide the higher level of confidence that a correct structure has been selected (Figures 3-1 & 3-2).

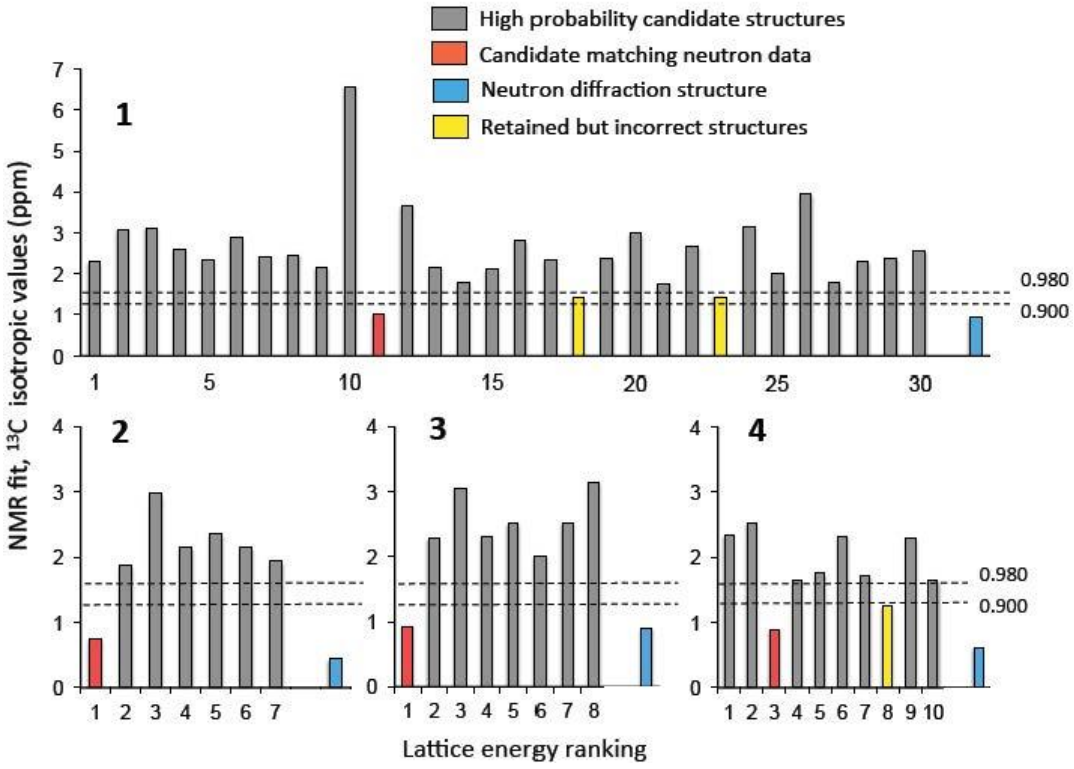


Figure 3-3: Agreement between experimental and computed isotropic ^{13}C shifts after refinement. Although the correct structure was selected as the best fit in all cases, three incorrect structures (one for **4** and two for **1**) could not be eliminated with high confidence. Here structures within 2 standard deviations of the global minimum were retained.

A notable feature of Figure 3-1 is the apparent improvement to the neutron diffraction coordinates of **4** due to the GIPAW refinement. Similar improvements were observed in the neutron coordinates of **1**–**3**. Neutron analyses can directly locate all atoms including hydrogens and are often considered the "gold standard" for structural determinations. Nevertheless, in all cases, GIPAW refinement of neutron coordinates improved NMR fits by a factor of 3–5 and decreased forces on the atoms by 2–3 orders of

magnitude as shown in Figure 3-4. These changes are particularly surprising given the extremely small changes in atom positions of less than 0.1 Å (all-atom root-mean-square deviation, RMSD). A visual comparison of the structures before and after refinement is given in Figure 3-5

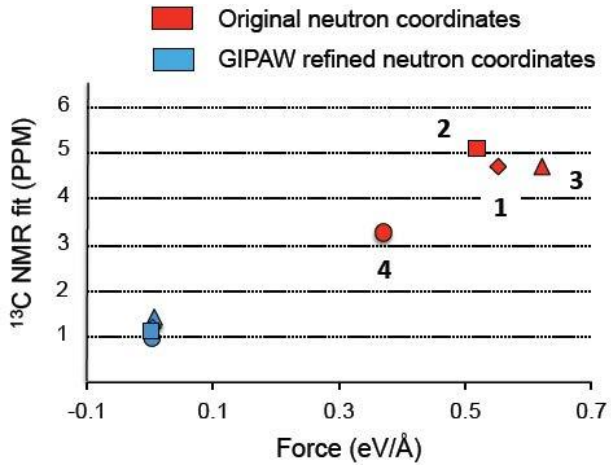


Figure 3-4: A plot showing improvement in neutron coordinates of structures **1**–**4** from a refinement that includes lattice fields. In all cases, NMR agreement improves by a factor of 3–5 and forces on the atoms decrease by 2–3 orders of magnitude. A comparison of atom positions before and after refinement (RMSD) was significantly less sensitive.

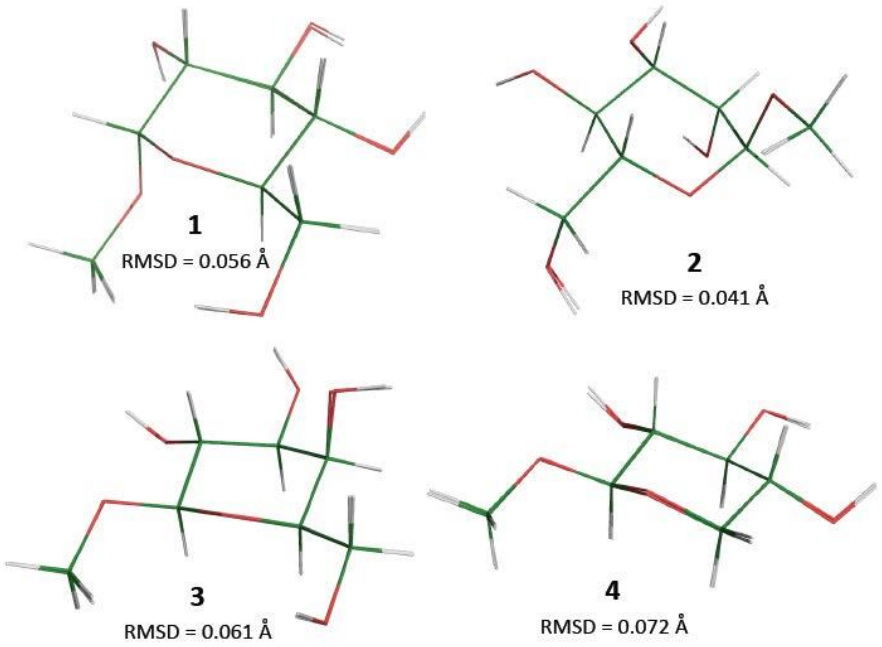


Figure 3-5: The original neutron coordinates are superimposed upon refined neutron coordinates to illustrate changes. In each case the structural differences are extremely small with the final coordinates found to lie within the error of the original neutron positions. Nevertheless, the structures differ with high statistical confidences (> 99.9%) based on improvements in NMR agreement and forces on the atoms.

We have previously reported several cases where single crystal neutron coordinates could be improved by refinements that included ^{13}C tensor values.[19,22] These refinements involved polycyclic aromatics, carbohydrates and a nucleoside. The present analysis extends these results by demonstrating improvement to four additional neutron structures and indicates that lattice-including refinement is creating highly accurate structures that rival single crystal neutron diffraction data. Further analysis involving an even greater variety of structures is needed to establish the general value of this approach. This study does not address errors in atom positions due to thermal movements. Inclusion of thermal factors will ultimately be needed to fully establish the accuracy of the structures obtained.

It is known that lattice-including refinements do not always converge to a global minimum,[12,22] thus the question remains whether that the final proposed coordinates could

be further refined. To test this possibility, the refined neutron and refined CSP coordinates for each structure were examined to see if they had converged to a common structure. The data in Table 3-2 demonstrate convergence in all cases in terms of NMR fit (tensor and isotropic data), forces upon the atoms and similarity in atomic positions. Convergence was assessed by statistically evaluating the differences in the final refined structures. The largest difference between two refined structures was for methyl β-D-xylopyranoside where the refined neutron structure was found to differ from the refined CSP structure at the 82% confidence level based on NMR tensor fit. Yet, even this difference was below the standard used (i.e. 95.4% or 2σ) to confidently state that two structures differed. Thus, in the present structures it appears that global minima have been identified in all cases.

Table 3-2: A comparison of several parameters to assess convergence of the refined neutron diffraction and the corresponding refined CSP structure to a common structure.

Structure	Source of data	NMR fit (ppm)[a] Tensor	Isotropic	Forces (eV Å^{-1})	RMSD (Å)[b]
Methyl β-D-xylopyranoside	Neutron diffraction	1.13	0.62	0.0016	–
	Theoretical CSP	1.41	0.89	0.0020	0.022
Methyl α-D-galactopyranoside	Neutron diffraction	1.43	0.89	0.0065	–
	Theoretical CSP	1.41	0.93	0.0066	0.037
Methyl α-D-mannopyranoside	Neutron diffraction	0.97	0.44	0.0030	–
	Theoretical CSP	1.13	0.74	0.0022	0.014
Methyl α-D-glucopyranoside	Neutron diffraction	1.34	0.97	0.0019	–
	Theoretical CSP	1.29	1.03	0.0018	0.086

[a]Values obtained by comparing principal values computed from GIPAW refined coordinates (PBE/Ultra-fine) to experimental principal values.
[b]A comparison of all atoms in the GIPAW refined structures to corresponding sites in the GIPAW refined neutron single crystal structures.

Conclusion

The structural selection and refinement process described herein represents a route to accurate crystal structure determinations solely from NMR data and computational methods. Neutron diffraction results are included only to verify the process. A requirement of this approach is that the CSP process must produce the correct structure among the numerous candidates. Currently, it appears that many prediction programs are capable of meeting this requirement based on results from blind studies.[24] Moreover, the results described here indicate that the lattice-including refinement process can refine initial structures that arise from different sources but which represent the same phase (e.g. x-ray single crystal and x-ray powder), into the same final structure. Thus, these results are expected to be somewhat independent of the specific CSP program used. Undoubtedly, the high accuracy of the ^{13}C tensor data used here ($\sigma = \pm 0.6$ ppm) helped the selection process. Errors in tensor principal value measurement methods in powders have been reported to range from ± 0.9 to ± 3 ppm[7d,25] for high quality data. This slightly larger error may decrease the selectivity for many powder samples and more work is needed to establish the impact of this difference on the selection process.

The present study focused only on carbohydrates and further study involving a more diverse set of structures is desirable in order to establish the general relevance of these methods. Currently, crystal structure predictions are infrequently used to solve actual structures because of ambiguity regarding how predicted structure correspond to experimental data. The process proposed here has the potential to resolve this ambiguity

and allow accurate structural analysis in materials that seldom form crystals or that are challenging to study for other reasons.

Acknowledgments

This work was supported by NSF grant CHE-0956755 to RJI. We thank Pennsylvania State University for providing access to the Lionxj metacluster (Research Computing and Cyperinfrastructure) for all GIPAW calculations. *The authors acknowledge the University of Central Florida Stokes Advanced Research Computing Center for providing computational resources and support that have contributed to results reported herein. URL: http://webstokes.ist.ucf.edu.*

Dedication

This work is dedicated to the memory of Prof. David M. Grant (1931–2013), graduate advisor to R. J. Iuliucci and J. K. Harper.

Supporting Information

Coordinates for refined neutron diffraction structures. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

[1]. Friedrich, W.; P. Knipping, P.; Laue, M. *Sitzungsber. Math. Phys. Phys. Kl. K. Bayer. Akad. Wiss. Munchen*, **1912**, pp 303–322.

[2]. Bragg, W. L. *Proc. R. Soc. London Ser. A* **1913**, *88*, 428-438.

[3]. Schwarzenbach, D. *Acta Crystallogr. Sect. A* **2012**, *68*, 57-67.

[4]. *Structure determination from powder diffraction data*; David, W. I. F., Shankland, K., McCusker, L. M., Baerlocher C. Eds,; Oxford: New York, **2002**.

[5]. Von Dreele, R. B. *J. Appl. Crystallogr.* **1999**, *32*, 1084-1089.

[6]. Basso, S.; Besnard, C.; Wright, J. P.; Margiolaki, I.; Fitch, A.; Pattison, P.; Schiltz, M. *Acta Cryst. Sect. D* **2010**, *66*, 756-761.

[7]. (a) Witter, R.; Sternburg, U.; Hesse, S.; Kondo, T.; Koch, F. T.; A. S. Ulrich, A. S. *Macromolecules*, **2006**, *39*, 6125-6132. (b) Brouwer, D. H. *J. Am. Chem. Soc.* **2008**, *130*, 6306-6307. (c) Harper, J. K.; Grant, D. M.; Zhang, Y.; Lee, P. L.; Von Dreele, R. B. *J. Am. Chem. Soc.* **2006**, *128*, 1547-1552. (d) Harper, J. K.; Doebbler, J. A.; Jacques, E.; Grant, D. M.; Von Dreele, R. B. *J. Am. Chem. Soc.* **2010**, *132*, 2928-2937. (e) Brouwer, D. H. *J. Magn. Reson.* **2009**, *194*, 136–146.

[8]. Harper, J. K.; Grant, D. M. *Cryst. Growth Des.* **2006**, *6*, 2315-2321.

[9]. Lehmann, C. W. *Angew. Chem. Int. Ed.* **2011**, *50*, 5616-5617.

[10]. Pickard, C. J.; Mauri, F. *Phys. Rev., B*, **2001**, *63*, 245101.

[11]. (a) Hutter, J.; Ballone, P.; Bernasconi, M.; Focher, P.; Fois, E.; Goedecker, S.; Parrinello, M.; Tuckerman, M. CPMD, 3.7 ed.; MPI für Festkörperforschung and IBM Zurich Research Laboratory, **1997-2001**; (b) Stueber, D. *Concepts Magn. Reson., Part A* **2006**, *28A*, 347-368.

[12]. Salager, E.; Day, D. M.; Stein, R. S.; Pickard, C. J.; Elena, B.; Emsley, L. *J. Am. Chem. Soc.* **2010**, *132*, 2564-2566.

[13]. Accelrys Software Inc., Materials Studio, Release 6.0, San Diego, **2011**.

[14]. Pfrommer, B. G.; Cote, M.; Louie, S. G.; Cohen, M. L. *J. Comput. Phys.*, **1997**, *131*, 233-240.

[15]. Liu, F.; Phung, C. G.; Alderman, D. W.; Grant, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 10629-10634.

[16]. Sherwood, M. H.; Alderman, D. W.; Grant, D. M. *J. Magn. Reson.* **1989**, *84*, 466-489.

[17]. (a) Takagi, S.; Jeffrey, G. A. *Acta Crystallogr., Sect. B* **1979**, *35*, 902-906. (b) Jeffrey, G. A.; McMullan, R. K.; Takagi, S. *Acta Crystallogr., Sect. B* **1977**, *33*, 728-737.

[18]. Alderman, D. W.; Sherwood, M. H.; Grant, D. M. *J. Magn. Reson. Ser. A* **1993**, *101*, 188-197.

[19]. Johnston, J. C.; Iuliucci, R. J.; Facelli, J. C.; Fitzgerald, G.; Mueller, K. T. *J. Chem. Phys.*, **2009**, *131*, 144503.

[20]. Olejniczak, S.; Ganicz, K.; Tomczykowa, M.; Gudej, J.; Potrzebowski, M. J. *J. Chem. Soc., Perkin 2*, **2002**, 1059-1065.

[21]. van Eijck, B. P.; Kroon, J. *J. Comput. Chem.* **1999**, *20*, 799-812.

[22]. (a) Harper, J. K.; Iuliucci, R. J.; Gruber, M.; Kalakewich, K. *CrystEngComm*, **2013**, *15*, 8693-8704. (b) Czernek, J.; Pawlak, T.; Potrzebowski, M. J.; Brus, J. *Chem. Phys. Lett.* **2013**, *555*, 135-140.

[23]. (a) Ashbrook, S. E.; Cutajar, M.; Pickard, C. J.; Walton R. I.; Wimperis, S. *Phys. Chem. Chem. Phys.*, **2008**, *10*, 5754-5764. (b) Profeta, M.; Mauri, F.; Pickard, C. J. *J. Am. Chem. Soc.*, **2003**, *125*, 541-548. (c) Brouwer, D. H.; Moudrakovski, I. L.; Darton, R. J.; R. E. Morris, R. E. *Magn. Reson. Chem.*, **2010**, *48*, S113-S121. (d) Pawlak, T.; Jaworska, M.; Potrzebowski, M. J. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3137-3145.

[24]. (a) D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J.

Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti and I. K. Zhitkov, *Acta Crystallogr. Sect. B*, 2011, **67**, 535-551. (b) G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Crystallogr. Sect. B*, 2009, **65**, 107-125. (c) G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt and P. Verwer, *Acta Crystallogr. Sect. B*, 2005, **61**, 511-527. (d) W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallogr., Sect B* **2002**, *58*, 647-661. (e) J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammin, J. D. Dunitz, A. Gavezzotti, D. W. M. Hoffmann, F. J. J. Leusen, W. T. M. Mooji, S. L. Price, B. Schweizer,

M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallogr., Sect. B*, 2000, **56**, 697-714.

[25]. Czernek, J.; Pawlak, T.; Potrzebowski, M. J. *Chem. Phys. Lett.* **2012**, *527*, 31-35.

CHAPTER 4: MONITORING THE REFINEMENT OF CRYSTAL STRUCTURES WITH ^{15}N SOLID-STATE NMR SHIFT TENSOR DATA

Reproduced from Kalakewich, K et al. Monitoring the Refinement of Crystal Structures with ^{15}N Solid-State NMR Shift Tensor Data. *J. Chem. Phys.* 2015, 143, 194702. With permission from AIP Publishing.

Keyton Kalakewich,[a] Robbie Iuliucci,[b] Karl T. Mueller,[c] Harriet Eloranta,[a] James K. Harper[a,*]

[a]Department of Chemistry, University of Central Florida, 4104 Libra Drive, Orlando, Florida, 32816, United States. [b]Department of Chemistry, Washington and Jefferson College, 60 Lincoln Street, Washington, Pennsylvania, 15301, United States. [c]Department of Chemistry, Penn State University, University Park, Pennsylvania, 16802, United States.

Abstract

The ^{15}N chemical shift tensor is shown to be extremely sensitive to lattice structure and a powerful metric for monitoring DFT refinements of crystal structures. These refinements include lattice effects and are applied here to five crystal structures. All structures improve based on a better agreement between experimental and calculated ^{15}N tensors, with an average improvement of 47.0 ppm. Structural improvement is further indicated by a decrease in forces on the atoms by 2–3 orders of magnitude and a greater similarity in atom positions to neutron diffraction structures. These refinements change bond lengths by more than the diffraction errors including adjustments to X–Y and X–H bonds (X, Y = C, N, and O) of 0.028 ± 0.002 Å and 0.144 ± 0.036 Å, respectively. The acquisition of ^{15}N tensors at natural abundance is challenging and this limitation is overcome by improved ^{1}H decoupling in the FIREMAT method. This decoupling dramatically narrows linewidths, improves signal-to-noise by up to 317%, and significantly improves the accuracy of measured tensors. A total of 39 tensors are measured with shifts distributed over a range of more than 400 ppm. Overall, experimental ^{15}N tensors are at

least 5 times more sensitive to crystal structure than ^{13}C tensors due to nitrogen's greater polarizability and larger range of chemical shifts.

Introduction

Over the past decade considerable emphasis has been placed on using solid-state NMR (SSNMR) data to gain insight into crystal structure. In early work SSNMR data were used to obtain structural information about a few atomic sites that were poorly defined in a diffraction structure. For example, the crystal structure of an STF type[1] zeolite, prepared to include fluoride, included an Si–F bond that was significantly longer than expected due to disorder.[2] A SSNMR analysis was less influenced by this disorder and therefore able to provide a more accurate Si–F distance of 1.744 Å.[3] This SSNMR information allowed the crystal structure to be re-refined and established that the environment around the silicon was trigonal bipyramidal rather than the tetrahedral structure suggested by x-ray diffraction. Solid-state NMR data have also been included in crystallographic studies to provide key distance measurements (e.g. $^{129}Xe/^{129}Xe$ in *t*-butylcalix[4]arene),[4] clarify the tautomeric form of a material,[5] or locate regions of disorder in otherwise highly ordered solids.[6] More recently, SSNMR methods have advanced to the point that structures of both individual molecules[7,8,9,10,11] and complete lattices[12,13,14] have now been determined without any type of diffraction data. Work in this area is now often referred to as *NMR crystallography* and a more complete description of the many advances in this area is given elsewhere.[15] Here, the primary focus is on one particular aspect of NMR crystallography, namely the use of solid-state NMR data to monitor DFT refinements of previously established crystal structures.

It has recently been demonstrated that such secondary DFT refinements performed using methods that include lattice effects can improve the structural quality of x-ray powder,[16,17,18,19] x-ray single crystal[19,20,21] and, remarkably, even neutron single crystal diffraction data.[14,19,22] The planewave DFT code CASTEP is the method usually used to include lattice effects in refinements. A recent advance, the efficient symmetry adapted clusters method,[23] offers a wider selection of functionals and could also be utilized in future studies. Often such refinements result in only minor changes to structure, but in some cases these adjustments provide new structural insights. For example, the original diffraction structure for cellulose Iα contained ambiguities at several O–H hydrogen orientations, suggesting that two significantly different hydrogen bonding arrangements were equally probable.[24] A lattice-including refinement unambiguously identified one of these arrangements as the only one consistent with experimental ^{13}C solid-state NMR evidence and demonstrated that dynamic interchange between these structures was improbable.[25] More recently, a CASTEP secondary refinement of the single-crystal x-ray diffraction structure of Lue-enkephalin clarified the bonding at two sites that were undetermined in the original structure.[26]

Structural improvement from a lattice-including refinement is indicated by better agreement between experimental NMR data and corresponding values computed from coordinates before and after refinement. Several types of NMR data are sensitive to this refinement including isotropic chemical shifts (for ^{27}Al, ^{29}Si, ^{13}C, ^{17}O, ^{15}N and ^{31}P nuclides),[12,14,16,17,20,27] quadrupolar coupling constants (for ^{17}O and ^{27}Al),[16,17] $^{1}H/^{1}H$ spin diffusion,[28] dipole couplings ($^{29}Si/^{29}Si$)[18,29] and chemical shift tensors (for ^{29}Si and ^{13}C).[14,18,20,22,29,30] In these studies, some types of data have been found to be more sensitive than others. The chemical shift tensor is particularly sensitive to structural refinement, and is

currently the only parameter that has been demonstrated to be sufficiently sensitive to monitor the further refinement of single crystal neutron diffraction data.[14,19,22] Up to this point in time, only shift tensors from ^{13}C and ^{29}Si have been used to monitor refinements. The focus here is to evaluate the influence of CASTEP refinements on ^{15}N tensor principal values. It is hypothesized that ^{15}N will be more sensitive to refinement than either ^{13}C or ^{29}Si since ^{15}N has a much larger shift range than either of these nuclides,[31] and nitrogen atoms often contains a polarizable lone electron pair that is extraordinarily sensitive to the local electronic environment. This expectation of greater sensitivity for ^{15}N is further supported by Duncan's extensive compilation of tensor principal values,[32] which includes ^{13}C, ^{15}N and ^{29}Si tensors. These data demonstrate that typical ^{13}C and ^{29}Si tensors from atoms in covalent bonds vary by approximately 300 and 320 ppm, respectively. In contrast, ^{15}N tensors in similar bonding environments vary by as much as 1830 ppm. Taken together, these considerations support the view that ^{15}N should be unusually sensitive to lattice refinement.

In this study, we introduce and exploit an improved method for rapidly obtaining accurate ^{15}N shift tensors from natural abundance samples. Tensors are measured from an unusually diverse range of ^{15}N environments, providing 39 data points distributed nearly uniformly in chemical shift space over a range of more than 400 ppm. The lattice-including refinements improve agreement between computed and experimental tensor data by nearly an order of magnitude. Most of the structural improvement is found to be in hydrogen positions, but non-hydrogen atom adjustments are also shown to contribute significantly. Calculating accurate ^{15}N tensors has often been problematic in prior studies. The data presented here demonstrate that very accurate ^{15}N tensors can now be computed when lattice effects are included in the NMR modeling and properly refined structures are utilized.

Experimental

Acetaminophen and L-histidine monohydrocholide monohydrate (histidine HCl H_2O) were purchased from Sigma-Aldrich with no elemental isotope enrichment and analyzed as received. Acetaminophen was evaluated by x-ray powder diffraction (as described elsewhere)[14] and found to be identical to the phase HXACAN26[33] deposited in the Cambridge Crystallographic Data Centre (CCDC). X-ray powder diffraction data were acquired for histidine HCl H_2O at room temperature using a Rigaku D/MAX diffractometer and Cu K alpha radiation (λ = 1.5406 Å). Data were collected from 2 - 40° using a step size of 0.04°. A comparison of the diffraction pattern with data previously reported from single crystal diffraction data indicated that the phase studied herein was HISTCM01.[34] Glycine and thymine were purchased as ^{15}N enriched samples (≥ 98 atom %) from Aldrich and evaluated by x-ray powder diffraction analysis using the parameters described above for histidine HCl H_2O. The glycine and thymine samples studied herein were found to be the same as GLYCIN18[35] and THYMIN01,[36] respectively. Cimetidine was obtained from Sigma-Aldrich and identified as form A (CIMETD03)[37] by the close match between the ^{13}C isotropic chemical shifts in the solid-state and those previously reported by Middleton et al.[38]

The ^{15}N FIREMAT data needed to measure tensors in for cimetidine were acquired on a CMX-400 Chemagnetics spectrometer operating at 40.549 MHz and employed SPINAL-64 ^{1}H decoupling[39] at 400.119 MHz. Evolution and acquisition dimension spectral widths were 10.8 and 30.3 kHz, respectively. Other parameters include a ^{1}H 90° pulse width of 3.9 μs, a ^{15}N 180 pulse width of 7.8 μs, and a cross polarization time of 6.0 ms using conventional (square) contact pulses on both channels. A total of 11 evolution increments of 3600 scans each were acquired together with a 13 s pulse delay for a total analysis time of 5.9 days. Data were rearranged to

extend the evolution dimension to 726 points according to a scheme described elsewhere.[40] A 7.5 mm PENCIL rotor module was employed and analysis was conducted at a spinning speed of 977 ± 1 Hz. The spectrum was externally referenced to the ^{15}N resonance of glycine at 33.4 ppm.[41]

The protonated and non-protonated ^{15}N sites in cimetidine were differentiated by performing a CP/MAS analysis that included a 50 μs period of interrupted ^{1}H decoupling before acquisition of the FID to eliminate all protonated resonances. All other acquisition parameters were identical to those described above for ^{15}N FIREMAT analysis.

The FIREMAT data acquired to measure tensors in histidine HCl H_2O were obtained using a 7.5 mm PENCIL probe with a CMX-200 Chemagnetics spectrometer operating at 20.276 MHz. SPINAL-64 ^{1}H decoupling was applied at a frequency of 200.044 MHz. Other parameters included a spinning speed of 330 ± 1 Hz, a cross polarization time of 3.0 ms, ^{1}H 90° and ^{15}N 180° pulse widths of 3.1 μs and 6.0 μs, respectively. Evolution and acquisition dimension spectral widths of 5.6 and 17.8 kHz were used, respectively. A total of 17 evolution increments of 1440 scans per increment were acquired with a pulse delay of 5 s. The spectrum was externally referenced to the ^{15}N resonance of glycine at 33.4 ppm. Data were rearranged to extend the evolution dimension to 645 points and processed utilizing procedures described elsewhere.[40,42]

All ^{15}N tensor data for acetaminophen, thymine and glycine were obtained from slow spinning 1D CP/MAS spectra with SPINAL-64 ^{1}H decoupling at 200.044 MHz. All data were acquired with a 7.5 mm PENCIL probe and a CMX-200 Chemagnetics spectrometer operating at 20.276 MHz with a cross polarization time of 3.0 ms. Spinning speeds of 1210, 360, and 150 Hz were used for analysis of acetaminophen, thymine and glycine, respectively. A spectral width of

5.6 kHz was utilized in all analyses and spectra were externally referenced to the ^{15}N resonance of glycine at 33.4 ppm. All tensor principal values reported herein were derived from experimental sideband data by a nonlinear least-squares fitting process using the banded matrix approach of Sethi *et al*. to compute sideband intensities.[42]

Computed chemical shielding values (σ) were converted to shift (δ) by fitting a plot of calculated shielding (after DFT refinement) versus experimental shift with a least-squares procedure. This process gave the equation $\sigma = -1.04 * \delta + 225.77$ with $R^2 = 0.992$ and a standard deviation in the computed σ of $\pm$ 7.0 ppm. Data were plotted after converting both experimental shifts and calculated shieldings into the icosahedral representation[43] where all points are equally weighted.

All calculations were performed using the Lion-XJ metacluster of the Research Computing and Cyperinfrastructure at Pennsylvania State University. The gauge including projector augmented wave (GIPAW)[44] NMR calculations were performed using the CASTEP module (Accelrys Software Inc.),[45] which inherently incorporates lattice factors into the calculations. Unit cell dimensions were not refined in any computations because these parameters are known to be systematically overestimated by the methods used.[16] The PBE functional was employed together with ultrasoft pseudopotentials. The so-called "ultra-fine" level was selected for the CASTEP parameters. The ultra-fine option uses planewave basis set cut-off energies of 390 eV for hydrocarbons and 550 eV if oxygen is present. The threshold convergence for SCF tolerance was 5×10^{-7} eV/atom. A k-point spacing of 0.071/Å was employed. The minimizing approach of Broyden, Fletcher, Goldfarb and Shanno[46] was used for geometry optimizations. For convergence of the optimization, a change in energy threshold of 5

x 10^{-6} eV/atom was reached, a maximum Cartesian force of 0.01 eV/Å on all atoms was achieved, and the maximum displacement of 5 x 10^{-4} Å for each atom was observed.

Results and Discussion

Nitrogen-15 chemical shift assignments

The five nitrogen-containing samples chosen for analysis appear in Figure 4-1. These compounds were selected with the aim of providing a wide range of nitrogen functional groups and include amine, aromatic, amide, imine, nitrile and imide moieties. Principal values of the chemical shift tensors obtained from these samples range from 4.0 to 410.3 ppm, providing the broad view of ^{15}N electronic environments required for a meaningful study. While each of the compounds selected has an established diffraction structure, several have obvious errors in hydrogen placement including significant errors in bond lengths (cimetidine) and unfeasible valence and torsion angles (cimetidine, glycine and acetaminophen). These structural shortcomings indicate the need for a secondary refinement.

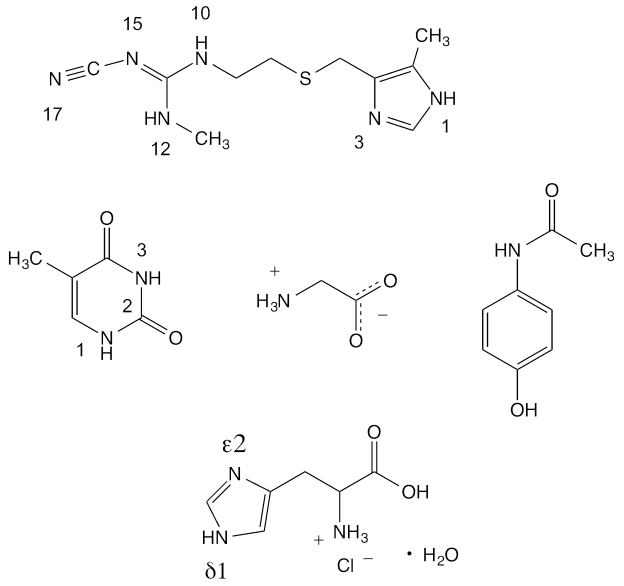


Figure 4-1: Structures of the compounds studied showing the nitrogen numbering. Structures evaluated include cimetidine (top), thymine, glycine, acetaminophen (middle, left to right, respectively) and histidine HCl H_2O (bottom).

Before accurate structural analysis can be performed chemical shifts must be correctly assigned to molecular positions. Two of the samples studied (glycine and acetaminophen) contain only one ^{15}N site and therefore do not require assignment. The ^{15}N shifts of histidine HCl H_2O have been previously assigned in the solid-state and those assignments are used herein.[47] For thymine, a comparison between experimental and computed shift tensor data was made to assign shifts. The assignment for thymine presented in Table 4-1 was preferred over the sole alternative with a statistical confidence of > 99.9% based on an F-test.[48] The ^{15}N shifts of cimetidine were initially sorted by experimentally differentiating the 3 protonated sites from the 3 non-protonated positions (see experimental). Within each sub-group, assignments were made by considering all permutations of experimental and computed tensor data and retaining the assignment

giving the best agreement. Assignments in cimetidine were made at > 99.9% confidence based on an F-test. All shift assignments are complied in Table 4-1.

Table 4-1: ^{15}N principal values and chemical shift assignments,[a] experimental (theoretical).

Compound	Position	δ_{11}	δ_{22}	δ_{33}	$\delta_{iso.}$
Cimetidine (form A)[b]	N1	248.2	176.2	86.5	170.3
		(266.1)	(152.9)	(70.7)	(163.2)
	N3[c]	312.2	252.9	4.0	189.7
		(306.8)	(241.2)	(-7.7)	(180.1)
	N10[c]	160.2	64.4	64.4	96.3
		(169.1)	(72.4)	(57.2)	(99.6)
	N12	157.7	58.3	33.3	83.1
		(156.6)	(61.2)	(43.2)	(87.0)
	N15[c]	129.3	81.3	46.0	85.5
		(126.6)	(75.9)	(45.3)	(82.6)
	N17	410.3	315.1	32.9	252.8
		(401.6)	(309.9)	(27.0)	(246.2)
Histidine HCl H_2O[b]	Nδ1	287.8	217.5	64.0	189.9
		(281.1)	(229.2)	(58.2)	(189.5)
	Nε2	276.6	195.1	57.8	176.5
		(270.8)	(203.7)	(52.4)	(175.6)
	NH_3^+	58.5	45.3	39.2	47.7
		(52.0)	(45.6)	(39.0)	(45.5)
Thymine[d]	N1	211.4	115.1	55.6	127.4
		(220.9)	(137.8)	(44.8)	(134.5)
	N3	225.8	146.9	98.5	157.1
		(235.8)	(155.1)	(87.3)	(159.4)
Glycine (γ-phase)[d]	N	42.3	34.3	23.7	33.4
		(39.3)	(32.1)	(22.9)	(31.4)
Acetaminophen[d]	N	240.5	85.4	85.3	137.1
		(262.1)	(91.3)	(71.4)	(141.6)

[a]All spectra were externally referenced to the ^{15}N peak in glycine at 33.4 ppm. [b]Data acquired using the FIREMAT experiment with SPINAL-64 decoupling. [c]Protonated positions based on dipolar-dephasing CP/MAS data (see experimental). [d]Data acquired from a 1D CP/MAS experiment with SPINAL-64 decoupling.

Modifying FIREMAT to improve ^{15}N tensor measurements

The ability to monitor crystal structure refinements with NMR is critically dependent upon the acquisition of accurate experimental NMR data. For measuring chemical shift tensors, the FIREMAT experiment has recently been shown to be one of the most accurate methods currently available.[49] FIREMAT is a slow spinning experiment (i.e. < 1 kHz typically) that displays the isotropic spectrum along one axis containing resonances for individual atomic sites and, in a second dimension, the spinning sideband patterns for each resonance that provide tensor values. Because of its high accuracy and ability to provide tensors for dozens of sites in a single experiment, FIREMAT was employed to obtain the majority of the ^{15}N tensors reported here.

The acquisition of ^{15}N tensor data is significantly more challenging than ^{13}C measurements since natural abundance ^{15}N is only 2.2% as sensitive as carbon. Isotopic enrichment is often used to overcome this insensitivity, but labeling is not always practical. To improve sensitivity and resolution in the FIREMAT experiment, a modified version of the sequence was employed that enhances sensitivity in natural abundance samples. The FIREMAT sequence was altered by replacing TPPM[50] (or CW) decoupling with the SPINAL-64 scheme.[39] It has previously been demonstrated that SPINAL decreases ^{15}N linewidths by as much as 44% (i.e. 12.2 to 6.8 Hz) under moderately fast spinning conditions of 13.3 kHz.[51] However, this improvement depends strongly on decoupling power and smaller differences are observed under most conditions. At higher spinning speeds (i.e. 33.33 kHz), both TPPM and SPINAL provided nearly identical ^{15}N linewidths and decoupling was significantly less effective.[51] However, SPINAL decoupling is preferred even at higher

spinning speeds because it is less influenced by experimental imperfections. For ^{13}C measurements, similar improvements in linewidths from SPINAL decoupling have been observed at slower spinning speeds of 5.7–8.8 kHz.[52] To our knowledge, no studies have evaluated SPINAL decoupling in solid state NMR at spinning rates less than 1 kHz, possibly because creation of a sideband free isotropic spectrum with correct peak amplitudes is challenging at this speed. Nevertheless, it is possible using, for example, the pseudo 2D spinning-sideband suppression technique[53] and the trend of decreasing line widths at lower spinning rates is an intriguing one, thus SPINAL-64 was incorporated in the FIREMAT experiment.

The FIREMAT experiment with SPINAL decoupling was evaluated using histidine HCl H_2O, a spinning rate of 330 Hz and a large 7.5 mm rotor to improve spinning stability. Sideband free isotropic spectra were obtained from these spectra using Gan's sideband suppression scheme.[53] Vastly improved linewidths were observed using SPINAL-64 when compared to a TPPM decoupled spectrum using identical analysis times, spinning speed and other parameters. Linewidths for all resonances are displayed on the spectra in Figure 4-2. In the most favorable case, linewidth decreased by a factor of 3.2 (i.e. 34.5 to 10.8 Hz) and the signal-to-noise ratio (S/N) increased by 317%. The FIREMAT spectra obtained are shown in Figure 4-3 and were acquired using the standard phase shifts originally suggested for SPINAL-64 and TPPM (see experimental).

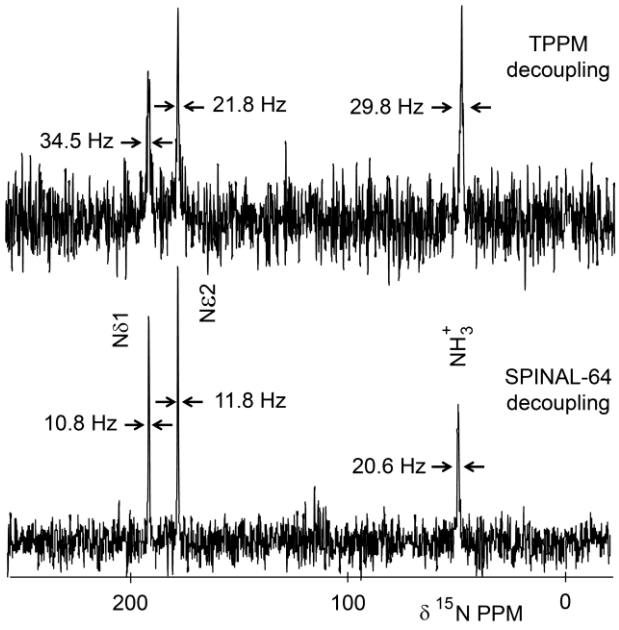


Figure 4-2: A comparison of ^{15}N linewidths in histidine HCl H_2O from TPPM versus SPINAL decoupling. In the most favorable case (i.e. Nδ1) SPINAL decoupling reduced linewidth by a factor of 3.2. On average, linewidths decreased to less than half their TPPM magnitudes due to SPINAL decouping. Both datasets were acquired using identical parameters and acquisition times and differ only in the type of of ^{1}H decoupling employed.

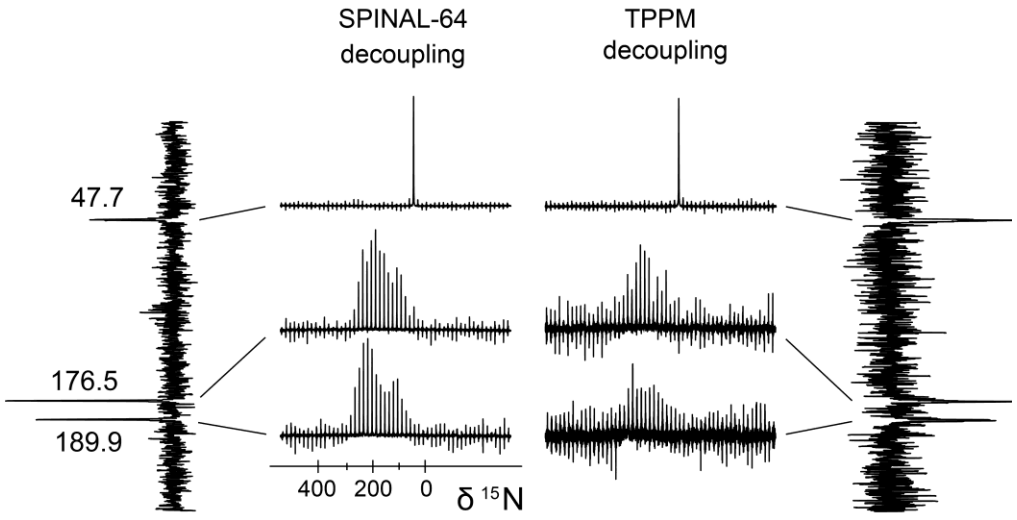


Figure 4-3. A comparison of the signal-to-noise (S/N) ratio in FIREMAT data acquired using SPINAL-64 (left) and TPPM decoupling. In each case, only the decoupling differed with all other parameters held constant. Improvements in S/N as large as 317% were observed from SPINAL decoupling. The accuracy of FIREMAT-SPINAL tensors also improved and is statistically indistinguishable from the accuracy of the corresponding single crystal tensor principal values.

One of the most significant observations from the FIREMAT-SPINAL data is that the ^{15}N tensor principal values obtained are considerably more accurate than those from FIREMAT-TPPM or, in fact, from several other tensor measurement techniques. Five groups have previously acquired ^{15}N tensor principal values for one or more of the nitrogens in histidine HCl H_2O and these data include NMR single crystal measurements. Single crystal data are usually regarded as the most accurate and provide a valuable reference point for comparisons. In the case of histidine HCl H_2O, the uncertainty in single crystal principal values is approximately $\pm$ 5.5 ppm based on linewidth (full-width at half-maximum) of the protonated imidazole nitrogen.[47] A comparison of the uncertainty in the ^{15}N tensors from FIREMAT-SPINAL and other methods versus single crystal values is given in Table 4-2. All uncertainties are calculated as root-mean-square deviations (RMSD). These data show that FIREMAT-SPINAL, with an RMSD of $\pm$ 2.8 ppm, provides tensors that lie within the uncertainty of single crystal values. In contrast, data from all other methods differs significantly from single crystal tensors.

Table 4-2: A comparison of ^{15}N tensor data for histidine HCl H_2O[a] from FIREMAT-SPINAL and several other methods showing the accuracy compared to the most accurately known data (bottom row).

Nδ			Nε			RMSD	
δ_{11}	δ_{22}	δ_{33}	δ_{11}	δ_{22}	δ_{33}	(ppm)	Ref.
287.8	217.5	64.0	276.6	195.1	57.8	± 2.8[b]	This study (SPINAL)
277.7	217.9	73.9	272.1	194.0	63.5	± 6.6[b]	[54]
272.0	226.2	69.7	268.4	198.4	61.8	± 7.7[c]	[55]
276.2	220.8	71.5	–	–	–	± 8.6[c]	[56]
273.3	218.3	77.1	269.3	194.7	59.7	± 9.6[b]	This study (TPPM)
260.4	203.4	77.4	–	–	–	± 24.3[c]	[57]
285.1	222.7	63.4	275.9	194.7	59.7	± 5.5[c,d]	47

[a]All analyses were conducted on the same phase, known in the Cambridge Crystallographic Database by the refcode HISTCM01.

[b,c]Referenced to either (*b*) an external glycine sample with the ^{15}N resonance assigned to 33.4 ppm relative to liquid NH_3 or (*c*) an external $^{15}NH_4Cl$ sample with the ^{15}N resonance assigned to 39.3 ppm relative to liquid NH_3.[41]

[d]These data represent the most accurately known ^{15}N tensor values for histidine HCl H_2O obtained using a large single crystal uniformly isotope labeled at > 98% ^{15}N. Error was estimated to be ± 5.5 ppm based on line width at half-height of the well-resolved resonances from the protonated imidazolium ring nitrogens.

The improved accuracy in the tensor data from SPINAL ^1H decoupling is likely due to the fact that slower spinning allows more decoupling cycles to be completed during a single rotor period. This process corresponds to greater averaging of the heteronuclear dipole coupling and superior accuracy in tensor measurements. It is noteworthy that many of the prior studies of histidine involved ^{15}N labeled samples with high S/N, thus the differences in accuracy between FIREMAT-SPINAL and other measurement methods are not due to differences in S/N. An exception is the SPINAL vs. TPPM comparison described herein where the TPPM decoupled spectrum (Figure 4-3) was obtained with low S/N in order to quantify sensitivity differences.

An illustration of the type of FIREMAT-SPINAL data obtained from a natural abundance sample that is more challenging than histidine is provided in Figure 4-4. These data for cimetidine require a long experiment time because several ^{15}N sites have large chemical shift anisotropies and the sample has a long ^{1}H spin-lattice relaxation time (T_1). Despite these obstacles, high S/N data were obtained utilizing 141.6 h of spectrometer time.

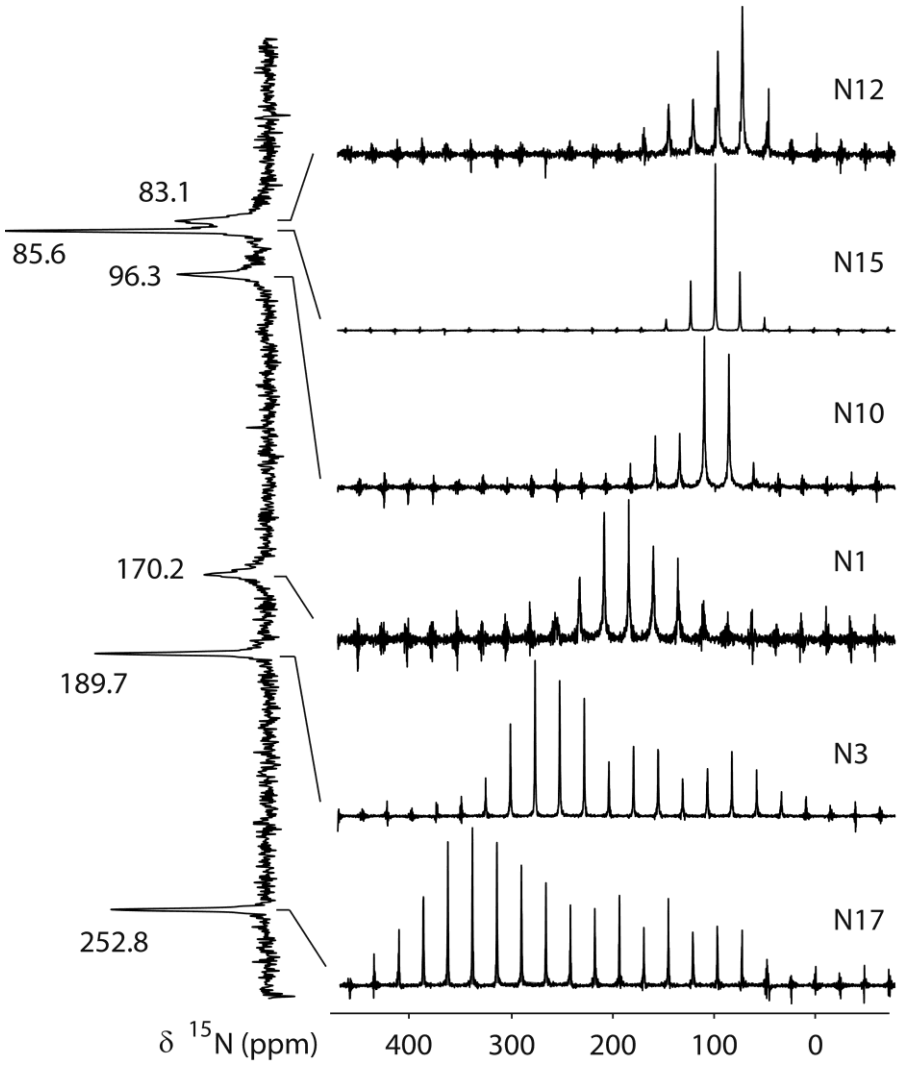


Figure 4-4: The ^{15}N FIREMAT spectrum of cimetidine, phase A, obtained using a natural abundance sample and SPINAL-64 ^{1}H decoupling.

Measuring the influence of lattice-including refinements on ^{15}N shift tensors

Pawlak and Potrzebowski have recently suggested that one can identify crystal structures likely to benefit from the lattice-including refinement by choosing structures that have R-values larger than 10%.[26] Their survey of data from the Cambridge Crystallographic Data Center found that approximately 22,000 structures are included in this category and could thus potentially be improved. Recent ^{13}C NMR work has demonstrated that many higher quality structures with R-values below 10 or even 5% can also benefit from CASTEP refinement.[14,19,20,26] These refinements usually involve minor adjustments in atom positions, but also serve to correct mistakes in diffraction data such as hydrogen atom misplacements,[19] ambiguities in hydrogen bonding arrangements[25] and the omission of certain bonds.[26] In the present study all structures evaluated have R-values ranging from 2.3 to 6.9% (Table 4-3), allowing further exploration of the influence of refinement on higher quality structures. Three of the structures studied here were obtained from x-ray single crystal diffraction. Structures for cimetidine and acetaminophen were determined using x-ray powder diffraction methods.

Table 4-3: The diffraction data type and R-values and for structures examined.

Structure	Data type	R-value (%)
Cimetidine, form A	X-ray powder	6.9
Thymine	X-ray single crystal	4.6
Glycine , γ-phase	X-ray single crystal	2.3
Acetaminophen	X-ray powder	Not given
Histamine HCl H_2O	X-ray single crystal	3.7

Tensor measurements were made without ^{15}N enrichment in three of the compounds (i.e. acetaminophen, cimetidine (phase A) and L-histidine HCl H_2O). These

three structures contributed 30 of the 39 principal values evaluated here. Two other products were purchased as uniformly ^{15}N labeled materials (glycine (γ-phase) and thymine). Experimental ^{15}N tensor principal values were obtained for all nitrogen sites (Table 4-1).

The correlation between calculated and experimental ^{15}N tensor data before and after crystal structure refinement is shown for all structures in Figure 4-5. Overall, the DFT refinement improves the RMS agreement[43] between experimental and calculated data from 54.0 ppm to 7.0 ppm when all atoms are adjusted. Changes to individual tensor values as large as 186.7 ppm (maximum absolute deviation) were observed. In contrast, monitoring comparable crystal structure refinements with ^{13}C principal values resulted in average changes of 1-8 ppm in most cases[19,20,21] with the largest change being roughly 62 ppm.[20]

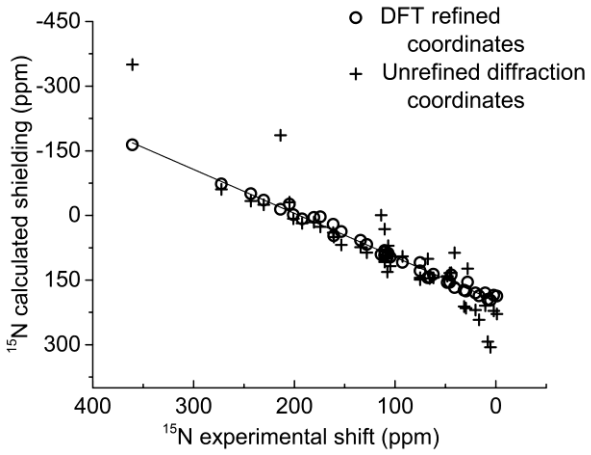


Figure 4-5: Plot of experimental ^{15}N principal values versus corresponding computed values before (red) and after (black) a lattice-including DFT refinement. Refinement significantly improves the R^2 from 0.833 to 0.992 and decreases the RMS uncertainty in computed shielding from 54.0 to 7.0 ppm. The best-fit slope and intercept for the data after refinement are, respectively, -1.037 and 225.771.

While these ^{15}N data suggest that more accurate coordinates are created by the lattice including refinement process, it is necessary to evaluate additional figures-of-merit to verify that all changes are consistent with structural improvements. Forces upon the atoms are provided at each step of the computation and were therefore also evaluated. In all cases, forces decreased by 2–3 orders of magnitude as a result of the refinement. A plot showing the improvements in the force and in the agreement between calculated and experimental NMR data for all structures is given in Figure 4-6.

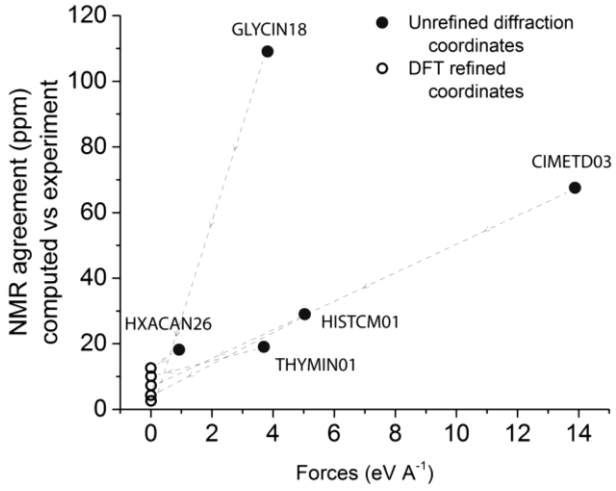


Figure 4-6: A comparison of the improvement in the RMS forces on the atoms of the lattice and in the agreement between computed and experimental ^{15}N shift tensors due to DFT refinement. Values calculated before and after refinement are shown, respectively, as black and unfilled circles. Individual structures are listed by refcodes as follows; GLYCIN18 = glycine, γ-phase, CIMETD03 = cimetidine, form A, HISTCM01 = histidine HCl H_2O, THYMIN01 = thymine, HXACAN26 = acetaminophen.

The data in 4-6 indicate that well refined structures have a high level of agreement between experimental and computed ^{15}N NMR tensor data and exhibit only small forces upon the atoms. Thus Pawlak and Potrzebowski's recommendation that crystal structures in need of secondary refinement can be identified by a high R-value can be expanded to include structures found to have large forces on the atoms in a lattice-including calculation

or that have a poor agreement between computed and experimental NMR tensor data. While the data in Figure 4-6 include only ^{15}N data, it has previously been found that ^{13}C tensor data show a similar trend and may also be used to assess crystal structure quality.[14]

Changes in atom positions and bond lengths from refinement

To further assess structural changes from the lattice-including refinements, a comparison was made of bond lengths before and after refinement. Prior work monitoring refinements with ^{13}C shift tensors revealed that the largest changes occur in bonds involving hydrogens,[14,19] presumably because hydrogen positions are poorly determined in single crystal x-ray studies. In this study, the largest changes in bond lengths were also found in bonds involving hydrogen, where an average change of ± 0.144 Å was observed. Changes to non-hydrogen atom bond lengths were smaller with average changes to C–C, C–N, C–O, C–S bond lengths of, respectively, ± 0.017 Å, ± 0.041 Å, ± 0.022 Å and ± 0.026 Å. The changes to bond lengths of individual structures are illustrated in 4-7. These changes are all significantly larger than the uncertainties in diffraction data for the corresponding bonds where values of ± 0.004 – 0.014 Å were reported for bonds involving non-hydrogen atoms and ± 0.04 – 0.07 Å for bonds involving hydrogen. A summary of bond length changes from CASTEP refinement and corresponding uncertainties in diffraction values are given in Table 4-4.

Table 4-4: Average changes in bond lengths from lattice-including DFT refinements (CASTEP) and the corresponding diffraction uncertainty.

Bond type	CASTEP changes (Å)	Diffraction error (Å, single crystal)	Diffraction error (Å, powder)[a]
C–C	± 0.017	± 0.004	± 0.015
C–N	± 0.041	± 0.004	± 0.015
C–O	± 0.022	± 0.004	–
C–S	± 0.026	–	± 0.011
C–H	± 0.14	± 0.04	± 0.10

[a]Crystal structures for acetaminophen and cimetidine were derived from x-ray powder diffraction data. All other structures were obtained from x-ray single crystal diffraction.

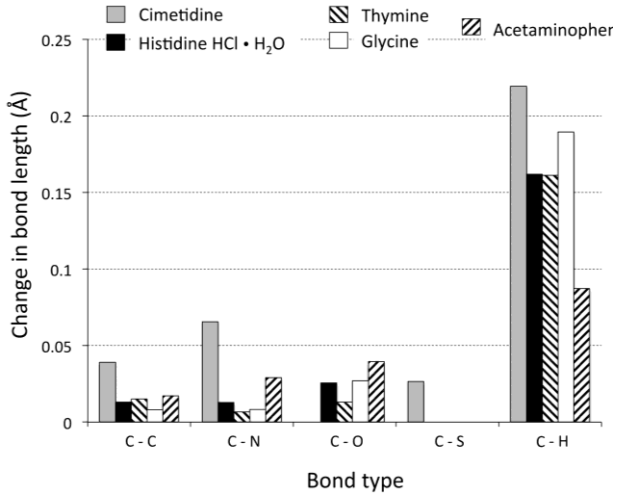


Figure 4-7: Changes in bond lengths resulting from the DFT lattice-including geometry refinement. In all cases structures were refined at the PBE/ultrafine level of theory. Some compounds lack certain bond types (e.g. C–S) and are therefore absent from the plot. The X–H bonds include contributions from C–H (18), N–H (9), and O–H (2).

It is notable that although these bond length adjustments are very strongly reflected in ^{15}N tensor data, they would be more difficult to detect using SSNMR methods that measure dipolar couplings. Many SSNMR structural studies now rely on distances obtained from dipole coupling methods and these techniques have uncertainties for typical structures of ± 0.1–0.7 Å for bonds between non-hydrogen atoms[58] and ± 0.04–0.07 Å for bonds involving hydrogens.[59] The changes from the lattice-including DFT refinements are thus at least an order of magnitude smaller than the detection limit of typical dipole coupling methods for non-hydrogen atoms. However, in many cases dipole coupling methods could measure changes to hydrogen positions.

Assessing structure quality

The structural changes from DFT refinements create molecules that differ statistically from the original diffraction coordinates at both hydrogen and non-hydrogen positions. An important question is whether these new structures represent legitimate improvements. While the superior NMR agreement and reduced forces on the atoms indicate improvement, the refined structures can be further evaluated by comparing atom positions with single crystal neutron diffraction coordinates for the same phase. Because single crystal neutron diffraction accurately locates both hydrogen and non-hydrogen atoms it is often regarded as the most accurate type of diffraction data. Such neutron data are available for acetaminophen, γ-glycine and histidine HCl H_2O, but not for the other structures, allowing a limited comparison. Atom positions were compared using the all-

atom root-mean-squared deviation (RMSD) in atom positions before and after refinement (Figure 4-8).

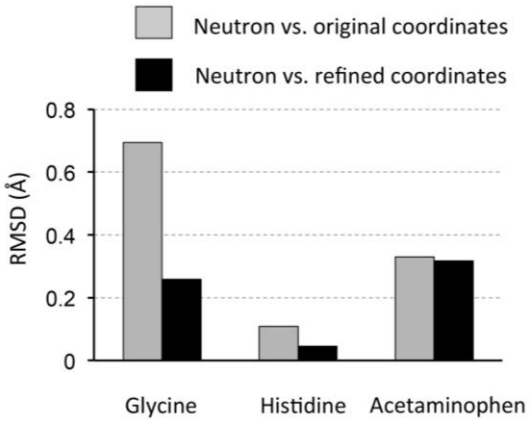


Figure 4-8: A comparison of RMSD in all atom positions for x-ray diffraction structures versus single crystal neutron diffraction coordinates of the same phase before and after refinement. Only acetaminophen, γ-glycine and histidine HCl H_2O are compared because these structures have both x-ray and single crystal neutron diffraction structures reported for the same phase.

The RMSD data shown in Figure 4-8 indicate that in the three cases considered here, the lattice-including refinement creates structures that more closely match single crystal neutron diffraction data than do the original coordinates. It is notable that acetaminophen shows only a very minor change upon refinement. However, a careful inspection of the refined structure revealed that the majority of the difference between this structure and the neutron coordinates arises from a difference in the position of the CH_3 hydrogens. However, the CH_3 hydrogen's are involved in a dynamic process involving rotation about the C–CH_3 axis. Such sites where dynamics alter atomic positions should not be included in the comparison of coordinates. If the CH_3 hydrogens are omitted, the RMSD before and after refinement are, respectively, 0.163 and 0.046 Å. Thus the DFT refinement process does, in fact, significantly improve the majority of the atomic positions in acetaminophen.

Overall, the metrics considered herein all support the conclusion that these DFT refinements are improving crystal structure.

Conclusions

This study describes the measurement of ^{15}N tensor principal values at natural abundance using the FIREMAT method with improved decoupling and explores the sensitivity of these tensor data to crystal structure refinement. Here, the CASTEP DFT method was used to further refine 5 crystal structures in an environment that includes lattice effects. Although structural changes were small in all cases with an all-atom RMSD of < 0.59 Å (see Supplemental material), the ^{15}N tensors are remarkably sensitive to the adjustments with changes as large as 186.7 ppm observed. On average the secondary refinement improved the agreement between computed and experimental ^{15}N tensors by 47.0 ppm. Forces on the atoms in the unit cell also decreased by 2–3 orders of magnitude and bond lengths changed by more than the uncertainty in the diffraction coordinates. A comparison of the refined coordinates versus neutron diffraction reference structures, where available, indicates that the changes from refinement represent genuine improvements. Comparing the results described herein to comparable ^{13}C tensor data[19,26,27] indicates that ^{15}N tensors are at least 5 times more sensitive to crystal structure refinement than are ^{13}C tensors. This improved sensitivity is most likely a reflection of nitrogen's greater chemical shift anisotropy.

Acknowledgments

This work was supported by National Science Foundation under CHE-1455159 To J.K.H. and CHE-0956755 to R.J.I. We acknowledge the University of Central Florida Stokes Advanced Research Computing Center for providing computational resources and support that have contributed to results reported herein, URL http://webstokes.ist.ucf.edu. We thank Pennsylvania State University for providing access to the Lionxj metacluster (Research Computing and Cyperinfrastructure) for all CASTEP calculations.

Supplemental Material

The DFT refined coordinates for all refined structures and a comparison of atom positions before and after CASTEP refinement.

References

[1]C. Baerlocher, L. B. McClusker, and D. H. Morris, *Atlas of Zeolite Framework Types*, (Elsevier, Amserdam, 2007) pp. 3–5.

[2]C. A. Fyfe, K. C. Wong-Moon, Y. Huang, and H. Grondey, *J. Am. Chem. Soc.* **117**, 10397 (1995).

[3]C. A. Fyfe, D. H. Brouwer, A. R. Lewis, L. A. Villaescusa, and R. E. Morris, *J. Am. Chem. Soc.* **124**, 7770 (2002).

[4]D. H. Brouwer, I. L. Moudrakovski, K. A. Udachin, G. D. Enright, and J. A. Ripmeester, *Cryst. Growth Des.* **8**, 1878 (2008).

[5]M. Tremayne, B. M. Kariuki, and K. D. M. Harris, *Angew. Chem. Intl. Engl. Ed.* **36**, 770 (1997).

[6]S. Meejoo, B. M. Kariuki, S. J. Kitchin, E. Y. Cheung, D. Albesa-Jové, and K. D. M. Harris, *Helv. Chim. Acta*, **86**, 1467 (2003).

[7]J. K. Harper, A. E. Mulgrew, J. Y. Li, D. H. Barich, G. A. Strobel, and D. M. Grant, *J. Am. Chem. Soc.* **123**, 9837 (2001).

[8]C. M. Rienstra, L. Tucker-Kellogg, C. P. Jaroniec, M. Hohwy, B. Reif, M. T. McMahon, B. Tidor, T. Lozano-Pérez, and R. G. Griffin, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10260 (2002).

[9]F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, and H. Oschkinat, *Nature* **420**, 98 (2002).

[10]J. K. Harper, D. H. Barich, J.-Z. Hu, G. A. Strobel, and D. M. Grant, *J. Org. Chem.* **68**, 4609 (2003).
[11]C. P. Jaroniec, C. E. MacPhee, V. S. Bajaj, M. T. McMahon, C. M. Dobson, and R. G. Griffin, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 711 (2004).
[12]D. H. Brouwer, S. Cadars, J. Eckert, Z. Li, O. Terasaki, and B. F. Chmelka, *J. Am. Chem. Soc.* **135**, 5641 (2013).
[13]M. Baias, J. –N. Dumez, P. H. Svensson, S. Schantz, G. M. Day, and L. Emsley, *J. Am. Chem. Soc.* **135**, 17501 (2013).
[14]K. Kalakewich, R. Iuliucci, and J. K. Harper, *Cryst. Growth Des.* **13**, 5391 (2013).
[15]R. K. Harris, R. E. Wasylishen, and M. J. Duer, *NMR Crystallography* (Wiley-Chichester, 2009).
[16]S. E. Ashbrook, M. Cutajar, C. J. Pickard, R. I. Walton, and S. Wimperis, *Phys. Chem. Chem. Phys.* **10**, 5754 (2008).
[17]M. Profeta, F. Mauri, and C. J. Pickard, *J. Am. Chem. Soc.* **125**, 541 (2003).
[18]D. H. Brouwer, *J. Am. Chem. Soc.* **130**, 6306 (2008).
[19]J. K. Harper, R. Iuliucci, M. Gruber, and K. Kalakewich, *CrystEngComm* **15**, 8693 (2013**)**.
[20]J. Czernek, T. Pawlek, M. J. Potrzebowski, and J. Brus, *Chem. Phys. Lett.* **555**, 135 (2013).
[21]T. Pawlak, M. Jaworska, and M. J. Potrzebowski, *Phys. Chem. Chem. Phys.* **15**, 3137 (2013).
[22]J. C. Johnston, R. J. Iuliucci, J. C. Facelli, G. Fitzgerald, and K. T. Mueller, *J. Chem. Phys.* **131**, 144503 (2009).
[23]S. T. Holmes, R. J. Iuliucci, K. T. Mueller, and C. Dybowski, C., *J. Chem. Phys.* **141**, 164121 (2014).
[24]Y. Nishiyama, J. Sugiyama, H. Chanzy, and P. Langan, *J. Am. Chem. Soc.* **125**, 14300 (2003).
[25]R. Witter, U. Sternberg, S. Hesse, T. Kondo, F.-T. Koch, and A. S. Ulrich, *Macromolecules* **39**, 6125 (2006).
[26]T. Pawlak and M. J. Potrzbowski, *J. Phys. Chem. B* **118**, 3298 (2014).
[27]R. A. Olsen, J. Struppe, D. W. Elliott, R. J. Thomas, and L. J. Mueller, *J. Am. Chem. Soc.* **125**, 11784 (2003).
[28]E. Salager, R. S. Stein, C. J. Pickard, B. Elena, and L. Emsley, *Phys. Chem. Chem. Phys.* **11**, 2610 (2009).
[29]D. H. Brouwer, *J. Magn. Reson.* **194**, 136 (2008).
[30]D. H. Brouwer, I. L. Moudrakovski, R. J. Darton, and R. E. Morris, *Magn. Reson. Chem.* **48**, S113 (2010).
[31]R. K. Harris and B. E. Mann, *NMR and the Periodic Table* (Academic Press, 1978).
[32]T. M. Duncan, *Principal Components of Chemical Shift Tensors: a Compilation*, (Farragut Press, Madison, 1997).
[33]M. Haisa, S. Kashino, R. Kawai, and H. Maeda, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **32**, 1283 (1976).
[34]K. Oda and H. Koyama, *Acta Crystallogr., Sect B: Crystallogr. Cryst. Chem.* **28**, 639 (1972).
[35]L. J. W. Shimon, M. Lohav, and L. Leiserowitz, *Nouv. J. Chim.* **10**, 723 (1986).
[36]G. Portalone, L. Bencivenni, M. Colapietro, A. Pieretti, and F. Ramondo, *Acta Chem. Scand.* **53**, 57 (1999).
[37]R. J. Cernik, A. K. Cheetham, C. K. Prout, D. J. Watkin, A. P. Wilkinson, and B. T. M. Willis, *J. Appl. Crystallogr.* **24**, 222 (1991).

[38]D. A. Middleton, D. C. S. Le, F. Berst, and D. G. Reid, *J. Pharm. Sci.* **86**, 1400 (1997).
[39]B. M. Fung, A. K. Khitrin, and K. Ermolaev, *J. Magn. Reson.* **142**, 97 (2000).
[40]D. W. Alderman, G. McGeorge, J. Z. Hu, R. J. Pugmire, and D. M. Grant, *Mol. Phys.* **95**, 1113 (1998).
[41]P. Bertani, J. Raya, and B. Bechinger, *Solid State Nucl. Magn. Reson.* **61-62**, 15 (2014).
[42]N. K. Sethi, D. W. Alderman, and D. M. Grant, *Mol. Phys.* **71**, 217 (1990).
[43]D. W. Alderman, M. H. Sherwood, and D. M. Grant, *J. Magn. Reson., Ser. A* **101**, 188 (1993).
[44]C. J. Pickard and F. Mauri, *Phys. Rev. B* **63**, 245101 (2001).
[45]Accelrys Software Inc., Materials Studio, Release 6.0 (San Diego, CA, 2011).
[46]B. G. Pfrommer, M. Cote, S. G. Louie, and M. L. Cohen, *J. Comput. Phys.* **131**, 233 (1997).
[47]G. Harbison, J. Herzfeld, and R. G. Griffin, *J. Am. Chem. Soc.* **103**, 4752 (1981).
[48]J. K. Harper and D. M. Grant, *J. Am. Chem. Soc.* **122**, 3708 (2000).
[49]J. K. Harper and R. J. Iuliucci, 2014. C-13 Chemical Shift Tensors in Organic Materials. *Encyclopedia of Analytical Chemistry*, Meyers, R. A. Ed., Wiley: New York, pp. 1-37.
[50]A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi, and R. G. Griffin, *J. Chem. Phys.* **103**, 6951 (1995).
[51]G. Comellas, J. J. Lopez, A. J. Nieuwkoop, L. R. Lemkau, and C. M. Rienstra, *J. Magn. Reson.* **209**, 131 (2011).
[52]R. S. Thakur, N. D. Kurur, and P. K. Madhu, *J. Magn. Reson.* **193**, 77 (2008).
[53]Z. Gan, *J. Magn. Reson., Ser. A* **109**, 253 (1994).
[54]M. Strohmeier, D. Stueber, and D. M. Grant, *J. Phys. Chem. A* **107**, 7629 (2003).
[55]Y. Wei, A. C. de Dios, and A. E. McDermott, *J. Am. Chem. Soc.* **121**, 10389 (1999).
[56]J. E. Roberts, G. S. Harbison, M. G. Munowitz, J. Herzfeld, and R. G. Griffin, *J. Am. Chem. Soc.* **109**, 4163 (1987).
[57]A. Ramamoorthy, C. H. Wu, and S. J. Opella, *J. Am. Chem. Soc.* **119**, 10479 (1997).
[58]J. M. Griffiths and R. G. Griffin, *Anal. Chim. Acta* **283**, 1081 (1993).
[59]B. –J. Van Rossum, C. P. de Groot, V. Ladizhansky, S. Vega, and H. J. M. de Groot, *J. Am. Chem. Soc.* **122**, 3465 (2000).

APPENDIX A: SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Contents.

The following is a compilation of calculated tensor values for GIPAW refined x-ray powder, x-ray single crystal and neutron single crystal diffraction structures. The PBE functional was employed together with the ultrasoft pseudopotential for all calculated values. Tensor values computed using unrefined neutron diffraction coordinates are included for comparison.

Table A-1 Acetaminophen SSNMR ^{13}C Principle Value Data

GIPAW refined neutron

Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1	218.97	136.78	35.51	130.42
C2	207.91	152.01	3.75	121.22
C3	195.96	129.31	13.20	112.82
C4	231.24	161.49	63.23	151.99
C5	194.29	128.30	16.70	113.10
C6	205.14	135.86	14.84	118.61
C7	236.70	171.01	88.31	165.34
C8	41.81	27.84	-4.69	21.65

GIPAW refined powder

Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1	219.20	135.68	35.56	130.15
C2	208.20	151.85	4.34	121.46
C3	196.37	129.48	13.82	113.22
C4	231.14	162.06	63.30	152.17
C5	195.22	129.84	16.29	113.78
C6	204.33	134.59	15.38	118.10
C7	238.41	169.27	87.95	165.21
C8	44.91	24.88	-9.89	19.97

	GIPAW refined X-ray single crystal			
Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1	218.96	136.91	35.70	130.52
C2	208.08	151.69	3.82	121.20
C3	195.98	129.41	13.25	112.88
C4	231.17	161.54	63.48	152.06
C5	194.45	128.56	16.72	113.24
C6	205.01	135.14	15.12	118.42
C7	237.05	170.76	88.66	165.49
C8	41.21	28.13	-4.78	21.52

	Neutron, no geometry optimization			
Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1	217.97	131.88	34.43	128.10
C2	206.08	150.45	0.13	118.89
C3	196.15	129.31	10.90	112.12
C4	226.55	154.34	61.47	147.45
C5	194.41	128.44	12.28	111.71
C6	207.07	134.14	10.18	117.13
C7	233.53	161.70	84.60	159.94
C8	30.60	14.33	-34.46	3.49

Table A-2 Adenosine SSNMR ^{13}C Principle Value Data

	GIPAW refined neutron			
Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C2	232.68	164.33	58.83	151.95
C4	214.95	161.18	57.83	144.65
C5	159.43	153.85	45.71	119.67
C6	201.76	193.71	52.73	149.40
C8	212.82	135.00	60.72	136.18
C1’	110.56	100.51	74.11	95.06
C2’	107.29	74.55	34.56	72.14
C3’	89.54	80.94	59.13	76.54
C4’	115.29	97.47	45.13	85.96
C5’	90.58	66.79	31.69	63.02

GIPAW refined powder

Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C2	232.93	162.99	59.73	151.88
C4	215.16	161.29	58.28	144.91
C5	159.32	154.01	45.89	119.74
C6	201.39	194.53	53.21	149.71
C8	212.89	134.30	60.90	136.03
C1’	110.96	101.17	73.41	95.18
C2’	107.43	74.87	34.73	72.34
C3’	90.06	81.96	58.98	77.00
C4’	115.99	97.58	46.04	86.54
C5’	91.44	66.95	31.60	63.33

GIPAW refined X-ray single crystal

Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C2	232.97	163.14	59.23	151.78
C4	215.15	161.21	58.10	144.82
C5	159.39	153.92	45.72	119.67
C6	201.49	194.22	52.92	149.54
C8	212.75	134.63	60.67	136.02
C1’	110.73	100.89	73.81	95.15
C2’	107.40	74.67	34.38	72.15
C3’	90.04	81.15	58.91	76.70
C4’	115.50	97.44	45.25	86.06
C5’	90.87	67.26	31.43	63.18

Neutron, no geometry optimization

Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C2	233.17	162.39	59.59	151.71
C4	214.47	160.99	55.63	143.70
C5	156.95	152.91	45.46	118.44
C6	199.48	194.71	51.14	148.44
C8	209.93	132.70	60.16	134.26
C1’	109.67	98.17	74.24	94.02
C2’	105.25	74.40	33.05	70.90
C3’	89.53	78.99	58.66	75.73
C4’	115.09	96.52	44.73	85.45
C5’	89.57	67.45	31.96	62.99

Table A-3 Naphthalene SSNMR ^{13}C Principle Value Data

	GIPAW refined neutron			
Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1, C5	224.04	143.53	17.37	128.32
C2, C6	226.82	139.74	3.98	123.51
C3, C7	226.49	139.71	3.53	123.24
C4, C8	222.72	149.10	14.28	128.70
C4a, C8a	205.39	201.99	-12.52	131.62

	GIPAW refined powder			
Position	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1, C5	224.20	143.29	17.50	128.33
C2, C6	226.97	139.66	4.06	123.56
C3, C7	226.77	139.33	3.82	123.31
C4, C8	222.73	148.50	14.55	128.59
C4a, C8a	205.28	201.88	-12.33	131.61

	GIPAW refined X-ray single crystal			
Positions	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1, C5	224.06	143.29	17.35	128.23
C2, C6	226.76	139.94	3.94	123.55
C3, C7	226.38	139.75	3.44	123.19
C4, C8	222.70	148.77	14.31	128.59
C4a, C8a	205.25	201.77	-12.53	131.50

	Neutron, no geometry optimization			
Positions	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)
C1, C5	223.84	142.23	15.58	127.22
C2, C6	225.71	138.21	0.39	121.44
C3, C7	224.93	138.17	0.41	121.17
C4, C8	221.54	146.68	11.28	126.50
C4a, C8a	203.50	200.46	-14.84	129.71

Force, RMDS, SSNMR and bond length data before and after GIPAW refinement.

The following is a compilation of calculated forces upon the atoms, the RMSD deviation of atomic positions and the agreement between calculated and experimental SSNMR ^{13}C shift tensor principal values before and after GIPAW refinement of the x-ray powder, x-ray single crystal and neutron single crystal diffraction structures. A comparison of the changes in bond lengths from refinement in also included as Table 9. All calculated values use the PBE functional together with the ultrasoft pseudopotential.

Powder data.

Table A-4: Comparison of the forces on the atoms in the unrefined powder (P) and the GIPAW refined powder (rP) structures.

	Forces (eV $Å^{-1}$)			
	P	rP (PBE)	rP (PW91)	rP (LDA)
Naphthalene	0.2515	0.0020	0.0030	0.0031
Acetaminophen	1.7435	0.0028	0.0029	0.0081
Adenosine	0.5361	0.0032	0.0047	0.0087

Table A-5: Comparison between experimental and computed ^{13}C tensor principal values for the unrefined powder (P) and GIPAW refined powder (rP) structures.

	SSNMR error (ppm)[a]			
	P	rP (PBE)	rP (PW91)	rP (LDA)
Naphthalene	4.25	3.01[b]	3.13	2.82
Acetaminophen	10.99	4.78[b]	4.80	4.95
Adenosine	4.50	4.16[b]	4.31	5.49

[a]All RMS error values are calculated by comparing experimental principal values to computed principal values obtained using the diffraction coordinates both before and after GIPAW refinement. All values were computed using Alderman's icosahedral representation (D. W. Alderman, G. McGeorge, J. Z. Hu, R. J. Pugmire and D. M. Grant, *Mol. Phys.*, 1998, **95**, 1113-1126).

[b]The SSNMR error for these refined structures differ statistically from errors computed using the original diffraction coordinates at statistical confidence levels of 90.00%, 99.99%, and 67.00%, respectively, for naphthalene, acetaminophen and adenosine.

X-ray single crystal data.

Table A-6: A comparison of the average forces on the atoms in the unrefined single crystal x-ray (SC) and the GIPAW refinement coordinates (rSC).

	Forces (eV Å^{-1})	
Structure	SC	rSC (PBE)
Naphthalene	1.3461	0.0045
Acetaminophen	3.8630	0.0021
Adenosine	3.2600	0.0027

Table A-7: Errors computed SSNMR ^{13}C tensor principal values for x-ray single crystal structures (SC) and the structures after GIPAW refinement (rSC).

	SSNMR error (ppm)[a]		
	SC	rSC (all atoms)	rSC (H only)
Naphthalene	5.34	3.11[b]	5.19
Acetaminophen	13.06	4.73[b]	5.81
Adenosine	9.65	4.13[b]	4.68

[a]All error values are obtained by comparing experimental principal values to computed principal values obtained using the diffraction coordinates both before and after GIPAW refinement. All values were computed using Alderman's icosahedral representation (D. W. Alderman, G. McGeorge, J. Z. Hu, R. J. Pugmire and D. M. Grant, *Mol. Phys.*, 1998, **95**, 1113-1126). The PBE functional was used for all calculations.
[b]The SSNMR error for these refined structures differ from errors computed using the original diffraction coordinates at a statistical confidence of 98%, 99.999% and 99.999%, respectively, for the naphthalene, acetaminophen and adenosine structures.

Table A-8: A comparison of RMSDs (Å) in atom positions for powder (P) and single crystal x-ray coordinates (SC) relative to the neutron diffraction values after refinement of all atoms and after refinement of only H atoms.

	RMSD (Å)					
	SC	rSC (all atoms)	rSC (H only)	P	rP (all atoms)	rP (H only)
Naphthalene	0.012	0.019[a]	-	0.013	0.019	0.013
Acetaminophen	0.067	0.053	0.053	0.163	0.046	0.039
Adenosine	0.093	0.086	0.035	0.354	0.109	0.064

[a]No hydrogen positions were reported in this structure, thus only carbon positions are compared.

Neutron single crystal data.

Table A-9: A comparison of the average forces on the atoms in the unrefined single crystal neutron diffraction structures (N) and these structures after GIPAW refinement (rN).

	Forces (eV Å^{-1})	
Structure	N	rN
Naphthalene	0.2210	0.0052
Acetaminophen	1.9608	0.0036
Adenosine	0.0753	0.0031

Table A-10: A comparison of experimental and computed SSNMR ^{13}C tensor principal values for the unrefined neutron single crystal (N) and GIPAW refined (rN) structures.

	SSNMR error (ppm)[a]	
	N	rN
Naphthalene	4.77	3.09[b]
Acetaminophen	10.05	4.75[b]
Adenosine	4.39	4.11[b]

[a]All error values are obtained by comparing experimental principal values to the corresponding computed principal values calculated using the diffraction coordinates both before and after GIPAW refinement. All values were computed using Alderman's icosahedral representation (D. W. Alderman, G. McGeorge, J. Z. Hu, R. J. Pugmire and D. M. Grant, *Mol. Phys.*, 1998, **95**, 1113-1126). The PBE functional was used for all calculations. [b]The SSNMR error for these refined structures can be said to differ from the pre-GIPAW structures at a statistical confidence level of 95.0%, 99.98% and 64.0%, respectively, for the naphthalene, acetaminophen and adenosine structures.

Table A-11: A comparison of the RMSDs (Å) in atomic positions of the unrefined single crystal neutron coordinates (N) versus GIPAW refined neutron (rN) and unrefined x-ray single crystal (SC) positions.

	RMSD (Å)[a]	
	N vs rN	N vs SC
Naphthalene	0.019	0.012
Acetaminophen	0.050	0.067
Adenosine	0.067	0.093

[a]These RMSDs compare all atoms in the structures.

A comparison of bond lengths from GIPAW refinement.

Table A-12: A comparison of differences (Å) in bond lengths for structures obtained from powder (P) or x-ray single crystal (SC) data versus neutron diffraction values (N) and the influence of GIPAW refinement on bond lengths.

Compound	Bond	P vs N[a]	rP vs N[a]	SC vs N	rSC vs N
Naphthalene	C – C	0.0059	0.0093	0.0119	0.0093
	C – H	H positions not reported in SC or P data			
Acetaminophen	C – C	0.0149	0.0209	0.0154	0.0213
	C – N	0.0184	0.0206	0.0032	0.0206
	C – O	0.0396	0.0282	0.0150	0.0300
	C – H	0.0639	0.0921	0.0986	0.0927
	N – H	0.0220	0.0520	0.0740	0.0510
	O – H	0.0360	0.0430	0.0790	0.0420
Adenosine	C – C	0.0112	0.0127	0.0076	0.0061
	C – N	0.0102	0.0072	0.0050	0.0059
	C – O	0.0151	0.0097	0.0049	0.0057
	C – H	NA	0.0101	0.1456	0.0041
	N – H	NA	0.0028	0.1643	0.0107
	O – H	NA	0.0135	0.1177	0.0173
Overall RMSD[b]	C – C	0.0109	0.0146	0.0116	0.0136
	C – N	0.0108	0.0072	0.0046	0.0059
	C – O	0.0203	0.0097	0.0073	0.0057
	C – H	0.0639	0.0101	0.1215	0.0036
	N – H	0.0220	0.0368	0.1274	0.0389
	O – H	0.0360	0.0135	0.1063	0.0233

[a]Differences in bond lengths for corresponding positions both before and after GIPAW refinement of the original diffraction data.
[b]The number of C–C, C–N, C–O bond included in this analysis was, respectively, 19, 12 and 7. The number of C–H, N–H, and O–H bonds evaluated was, respectively, 15, 3 and 4.

GIPAW refined crystal structure coordinates.

This file contains a total of nine sets of refined diffraction coordinates in CIF format. The structures include the refined x-ray powder, x-ray single crystal and neutron single crystal coordinates for adenosine, acetaminophen and naphthalene. All structures were refined at the GGA-PBE level of theory using parameters summarized in the Experimental section of the manuscript. In all cases, the reported diffraction coordinates were used as an initial starting point.

```
data_ADENOS01_neutron
_audit_creation_date              2012-05-17
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21'
_symmetry_Int_Tables_number       4
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x,y+1/2,-z
_cell_length_a                    4.7885
_cell_length_b                    10.2400
_cell_length_c                    11.7720
_cell_angle_alpha                 90.0000
_cell_angle_beta                  99.5900
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
N1     N     0.91604  -0.09576   0.18337   0.00000  Uiso   1.00
C1     C     0.79057  -0.07928   0.27588   0.00000  Uiso   1.00
N2     N     0.63963   0.02347   0.30156   0.00000  Uiso   1.00
C2     C     0.62382   0.11669   0.21963   0.00000  Uiso   1.00
C3     C     0.75155   0.11459   0.12069   0.00000  Uiso   1.00
C4     C     0.90613   0.00035   0.10263   0.00000  Uiso   1.00
N3     N     0.68784   0.22774   0.05719   0.00000  Uiso   1.00
C5     C     0.52518   0.29516   0.11643   0.00000  Uiso   1.00
N4     N     0.48000   0.23337   0.21530   0.00000  Uiso   1.00
N5     N     1.03718  -0.01899   0.01237   0.00000  Uiso   1.00
C6     C     0.31467   0.27816   0.30229   0.00000  Uiso   1.00
C7     C     0.50337   0.31456   0.41675   0.00000  Uiso   1.00
C8     C     0.54301   0.46147   0.40083   0.00000  Uiso   1.00
C9     C     0.25790   0.50394   0.33088   0.00000  Uiso   1.00
C10    C     0.27610   0.61889   0.25152   0.00000  Uiso   1.00
```

```
O1     O     0.15655   0.39012   0.26222   0.00000  Uiso   1.00
O2     O     0.35410   0.29253   0.50863   0.00000  Uiso   1.00
O3     O     0.63364   0.53214   0.50338   0.00000  Uiso   1.00
O4     O     0.00591   0.64375   0.18404   0.00000  Uiso   1.00
H1     H     0.81257  -0.15891   0.33790   0.00000  Uiso   1.00
H2     H     0.43118   0.38960   0.09262   0.00000  Uiso   1.00
H3     H     1.12390  -0.10935   0.00266   0.00000  Uiso   1.00
H4     H     1.02592   0.04842  -0.05413   0.00000  Uiso   1.00
H5     H     0.17022   0.19808   0.31561   0.00000  Uiso   1.00
H6     H     0.70706   0.26198   0.42912   0.00000  Uiso   1.00
H7     H     0.70693   0.47429   0.34644   0.00000  Uiso   1.00
H8     H     0.10765   0.52556   0.39000   0.00000  Uiso   1.00
H9     H     0.35018   0.70473   0.30483   0.00000  Uiso   1.00
H10    H     0.43620   0.59879   0.19601   0.00000  Uiso   1.00
H11    H     0.34154   0.19535   0.51399   0.00000  Uiso   1.00
H12    H     0.51677   0.51316   0.56514   0.00000  Uiso   1.00
H13    H    -0.01493   0.74147   0.17365   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
_ccdc_geom_bond_type
N1     C1      1.339   .     S
N1     C4      1.364   .     S
C1     N2      1.339   .     S
C1     H1      1.088   .     S
N2     C2      1.350   .     S
C2     C3      1.403   .     S
C2     N4      1.376   .     S
C3     C4      1.419   .     S
C3     N3      1.385   .     S
C4     N5      1.336   .     S
N3     C5      1.323   .     S
C5     N4      1.373   .     S
C5     H2      1.084   .     S
N4     C6      1.468   .     S
N5     H3      1.028   .     S
N5     H4      1.038   .     S
C6     C7      1.538   .     S
C6     O1      1.411   .     S
C6     H5      1.101   .     S
C7     C8      1.532   .     S
C7     O2      1.410   .     S
C7     H6      1.102   .     S
C8     C9      1.535   .     S
C8     O3      1.412   .     S
C8     H7      1.100   .     S
C9     C10     1.514   .     S
C9     O1      1.455   .     S
C9     H8      1.104   .     S
C10    O4      1.424   .     S
```

```
C10    H9      1.104   .     S
C10    H10     1.106   .     S
O2     H11     1.000   .     S
O3     H12     1.008   .     S
O4     H13     1.011   .     S
#===END


data_ADENOS10_x-ray
_audit_creation_date              2012-05-23
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21'
_symmetry_Int_Tables_number       4
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x,y+1/2,-z
_cell_length_a                    4.8250
_cell_length_b                    10.2820
_cell_length_c                    11.8230
_cell_angle_alpha                 90.0000
_cell_angle_beta                  99.3000
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1     C    -0.68070   0.28035   0.30125   0.00000  Uiso   1.00
C2     C    -0.21590  -0.07813   0.27471   0.00000  Uiso   1.00
C3     C    -0.49428   0.31797   0.41475   0.00000  Uiso   1.00
C4     C    -0.45989   0.46471   0.39871   0.00000  Uiso   1.00
C5     C    -0.37531   0.11806   0.21927   0.00000  Uiso   1.00
C6     C    -0.74386   0.50492   0.32926   0.00000  Uiso   1.00
C7     C    -0.24667   0.11538   0.12099   0.00000  Uiso   1.00
C8     C    -0.72965   0.62038   0.25078   0.00000  Uiso   1.00
C9     C    -0.09638   0.00066   0.10267   0.00000  Uiso   1.00
C10    C    -0.46773   0.29621   0.11663   0.00000  Uiso   1.00
H1     H    -0.19730  -0.15767   0.33621   0.00000  Uiso   1.00
H2     H    -0.89457   0.52432   0.38818   0.00000  Uiso   1.00
H3     H    -0.65881   0.70587   0.30419   0.00000  Uiso   1.00
H4     H    -0.57017   0.60196   0.19513   0.00000  Uiso   1.00
H5     H    -1.02046   0.74117   0.17342   0.00000  Uiso   1.00
H6     H    -0.55919   0.39056   0.09277   0.00000  Uiso   1.00
H7     H     0.11768  -0.10963   0.00259   0.00000  Uiso   1.00
H8     H     0.02512   0.04797  -0.05329   0.00000  Uiso   1.00
H9     H    -0.82314   0.20042   0.31529   0.00000  Uiso   1.00
```

```
H10    H    -0.29075   0.26689   0.42621   0.00000  Uiso   1.00
H11    H    -0.64911   0.19841   0.51347   0.00000  Uiso   1.00
H12    H    -0.29732   0.47923   0.34446   0.00000  Uiso   1.00
H13    H    -0.48643   0.51431   0.56273   0.00000  Uiso   1.00
N1     N    -0.09024  -0.09527   0.18270   0.00000  Uiso   1.00
N2     N     0.03366  -0.01926   0.01259   0.00000  Uiso   1.00
N3     N    -0.36334   0.02503   0.30052   0.00000  Uiso   1.00
N4     N    -0.30637   0.22832   0.05788   0.00000  Uiso   1.00
N5     N    -0.51578   0.23494   0.21487   0.00000  Uiso   1.00
O1     O    -0.83823   0.39126   0.26043   0.00000  Uiso   1.00
O2     O    -0.63997   0.29520   0.50711   0.00000  Uiso   1.00
O3     O    -0.37332   0.53541   0.50079   0.00000  Uiso   1.00
O4     O    -0.99817   0.64405   0.18409   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
_ccdc_geom_bond_type
C1     C3      1.539   .     S
C1     H9      1.101   .     S
C1     N5      1.469   .     S
C1     O1      1.412   .     S
C2     H1      1.088   .     S
C2     N1      1.340   .     S
C2     N3      1.340   .     S
C3     C4      1.533   .     S
C3     H10     1.102   .     S
C3     O2      1.411   .     S
C4     C6      1.536   .     S
C4     H12     1.101   .     S
C4     O3      1.413   .     S
C5     C7      1.403   .     S
C5     N3      1.350   .     S
C5     N5      1.377   .     S
C6     C8      1.515   .     S
C6     H2      1.104   .     S
C6     O1      1.454   .     S
C7     C9      1.419   .     S
C7     N4      1.385   .     S
C8     H3      1.104   .     S
C8     H4      1.107   .     S
C8     O4      1.425   .     S
C9     N1      1.364   .     S
C9     N2      1.336   .     S
C10    H6      1.084   .     S
C10    N4      1.323   .     S
C10    N5      1.374   .     S
H5     O4      1.010   .     S
H7     N2      1.028   .     S
H8     N2      1.037   .     S
H11    O2      1.000   .     S
```

```
H13    O3      1.005   .     S
#===END


data_ADENOS11_powder_H_added
_audit_creation_date              2012-05-17
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21'
_symmetry_Int_Tables_number       4
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x,y+1/2,-z
_cell_length_a                    4.8392
_cell_length_b                    10.2908
_cell_length_c                    11.8569
_cell_angle_alpha                 90.0000
_cell_angle_beta                  99.3090
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
N1     N     0.90054  -0.10255   0.18321   0.00000  Uiso   1.00
C2     C     0.77575  -0.08455   0.27499   0.00000  Uiso   1.00
N3     N     0.63252   0.01959   0.30113   0.00000  Uiso   1.00
C4     C     0.62394   0.11268   0.22040   0.00000  Uiso   1.00
C5     C     0.75201   0.10913   0.12238   0.00000  Uiso   1.00
C6     C     0.89803  -0.00651   0.10378   0.00000  Uiso   1.00
N7     N     0.69662   0.22254   0.05986   0.00000  Uiso   1.00
C8     C     0.53822   0.29150   0.11869   0.00000  Uiso   1.00
N9     N     0.48783   0.23043   0.21635   0.00000  Uiso   1.00
N10    N     1.02780  -0.02699   0.01406   0.00000  Uiso   1.00
C11    C     0.32532   0.27669   0.30275   0.00000  Uiso   1.00
C12    C     0.51322   0.31652   0.41482   0.00000  Uiso   1.00
C13    C     0.54249   0.46325   0.39759   0.00000  Uiso   1.00
C14    C     0.25760   0.50112   0.32835   0.00000  Uiso   1.00
C15    C     0.26804   0.61537   0.24849   0.00000  Uiso   1.00
O16    O     0.16500   0.38628   0.26130   0.00000  Uiso   1.00
O17    O     0.37196   0.29402   0.50807   0.00000  Uiso   1.00
O18    O     0.62597   0.53482   0.49914   0.00000  Uiso   1.00
O19    O    -0.00064   0.63720   0.18211   0.00000  Uiso   1.00
H20    H     0.79151  -0.16421   0.33602   0.00000  Uiso   1.00
H21    H     0.45044   0.38646   0.09523   0.00000  Uiso   1.00
H22    H     1.02257   0.04065  -0.05102   0.00000  Uiso   1.00
H23    H     1.11076  -0.11740   0.00385   0.00000  Uiso   1.00
```

```
H24    H     0.18574   0.19650   0.31845   0.00000  Uiso   1.00
H25    H     0.71757   0.26671   0.42569   0.00000  Uiso   1.00
H26    H     0.70405   0.47883   0.34345   0.00000  Uiso   1.00
H27    H     0.10819   0.52087   0.38739   0.00000  Uiso   1.00
H28    H     0.42652   0.59660   0.19284   0.00000  Uiso   1.00
H29    H     0.33761   0.70184   0.30050   0.00000  Uiso   1.00
H30    H     0.36116   0.19734   0.51447   0.00000  Uiso   1.00
H31    H     0.51567   0.51179   0.56123   0.00000  Uiso   1.00
H32    H    -0.02771   0.73420   0.17290   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
_ccdc_geom_bond_type
N1     C2      1.340   .     S
N1     C6      1.364   .     S
C2     N3      1.340   .     S
C2     H20     1.088   .     S
N3     C4      1.350   .     S
C4     C5      1.403   .     S
C4     N9      1.376   .     S
C5     C6      1.419   .     S
C5     N7      1.385   .     S
C6     N10     1.337   .     S
N7     C8      1.323   .     S
C8     N9      1.374   .     S
C8     H21     1.084   .     S
N9     C11     1.468   .     S
N10    H22     1.036   .     S
N10    H23     1.028   .     S
C11    C12     1.539   .     S
C11    O16     1.411   .     S
C11    H24     1.101   .     S
C12    C13     1.533   .     S
C12    O17     1.410   .     S
C12    H25     1.103   .     S
C13    C14     1.536   .     S
C13    O18     1.414   .     S
C13    H26     1.101   .     S
C14    C15     1.516   .     S
C14    O16     1.454   .     S
C14    H27     1.104   .     S
C15    O19     1.425   .     S
C15    H28     1.107   .     S
C15    H29     1.104   .     S
O17    H30     1.000   .     S
O18    H31     1.005   .     S
O19    H32     1.010   .     S
#===END
```

```
data_HXACAN01_x-ray
_audit_creation_date              2012-05-23
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21/A'
_symmetry_Int_Tables_number       14
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x+1/2,y+1/2,-z
  -x,-y,-z
  x+1/2,-y+1/2,z
_cell_length_a                    12.9300
_cell_length_b                    9.4000
_cell_length_c                    7.1000
_cell_angle_alpha                 90.0000
_cell_angle_beta                  115.9000
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1     C    -0.06411   0.35155  -0.15085   0.00000  Uiso   1.00
C2     C     0.00863   0.34873  -0.25040   0.00000  Uiso   1.00
C3     C     0.10399   0.25909  -0.17703   0.00000  Uiso   1.00
C4     C     0.12882   0.17090  -0.00355   0.00000  Uiso   1.00
C5     C     0.05727   0.17571   0.09828   0.00000  Uiso   1.00
C6     C    -0.03781   0.26520   0.02531   0.00000  Uiso   1.00
C7     C    -0.22220   0.50524  -0.39935   0.00000  Uiso   1.00
C8     C    -0.32404   0.59084  -0.41443   0.00000  Uiso   1.00
H1     H    -0.00915   0.41624  -0.38510   0.00000  Uiso   1.00
H2     H     0.16055   0.25967  -0.25449   0.00000  Uiso   1.00
H3     H     0.07579   0.10846   0.23413   0.00000  Uiso   1.00
H4     H    -0.09310   0.26841   0.10537   0.00000  Uiso   1.00
H5     H     0.24275   0.04200   0.20291   0.00000  Uiso   1.00
H6     H    -0.19661   0.44554  -0.10771   0.00000  Uiso   1.00
H7     H    -0.39822   0.57090  -0.56505   0.00000  Uiso   1.00
H8     H    -0.34716   0.56898  -0.28595   0.00000  Uiso   1.00
H9     H    -0.30394   0.70461  -0.41211   0.00000  Uiso   1.00
N1     N    -0.16249   0.43862  -0.21400   0.00000  Uiso   1.00
O1     O     0.22201   0.08291   0.06000   0.00000  Uiso   1.00
O2     O    -0.19471   0.49883  -0.54792   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
```

```
_ccdc_geom_bond_type
C1     C2      1.402   .     S
C1     C6      1.403   .     S
C1     N1      1.411   .     S
C2     C3      1.393   .     S
C2     H1      1.085   .     S
C3     C4      1.401   .     S
C3     H2      1.090   .     S
C4     C5      1.402   .     S
C4     O1      1.366   .     S
C5     C6      1.390   .     S
C5     H3      1.089   .     S
C6     H4      1.091   .     S
C7     C8      1.507   .     S
C7     N1      1.354   .     S
C7     O2      1.253   .     S
C8     H7      1.097   .     S
C8     H8      1.096   .     S
C8     H9      1.099   .     S
H5     O1      1.006   .     S
H6     N1      1.029   .     S
#===END


data_HXACAN19_neutron
_audit_creation_date              2012-05-17
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21/A'
_symmetry_Int_Tables_number       14
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x+1/2,y+1/2,-z
  -x,-y,-z
  x+1/2,-y+1/2,z
_cell_length_a                    12.8720
_cell_length_b                    9.3700
_cell_length_c                    7.0850
_cell_angle_alpha                 90.0000
_cell_angle_beta                  115.6200
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1     C    -0.06547   0.35300  -0.15123   0.00000  Uiso   1.00
```

```
C2     C     0.00815   0.34995  -0.25012   0.00000  Uiso   1.00
C3     C     0.10346   0.25968  -0.17684   0.00000  Uiso   1.00
C4     C     0.12752   0.17099  -0.00364   0.00000  Uiso   1.00
C5     C     0.05536   0.17643   0.09822   0.00000  Uiso   1.00
C6     C    -0.03977   0.26647   0.02495   0.00000  Uiso   1.00
C7     C    -0.22388   0.50611  -0.40072   0.00000  Uiso   1.00
C8     C    -0.32647   0.59154  -0.41639   0.00000  Uiso   1.00
N1     N    -0.16423   0.44015  -0.21497   0.00000  Uiso   1.00
O1     O     0.22040   0.08204   0.05939   0.00000  Uiso   1.00
O2     O    -0.19584   0.49928  -0.54913   0.00000  Uiso   1.00
H1     H    -0.00911   0.41771  -0.38458   0.00000  Uiso   1.00
H2     H     0.16062   0.26006  -0.25395   0.00000  Uiso   1.00
H3     H     0.07343   0.10906   0.23410   0.00000  Uiso   1.00
H4     H    -0.09562   0.27000   0.10484   0.00000  Uiso   1.00
H5     H     0.24103   0.04110   0.20256   0.00000  Uiso   1.00
H6     H    -0.19880   0.44722  -0.10878   0.00000  Uiso   1.00
H7     H    -0.39937   0.57448  -0.56937   0.00000  Uiso   1.00
H8     H    -0.35212   0.56655  -0.29123   0.00000  Uiso   1.00
H9     H    -0.30532   0.70578  -0.40789   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
_ccdc_geom_bond_type
C1     C2      1.401   .     S
C1     C6      1.403   .     S
C1     N1      1.411   .     S
C2     C3      1.393   .     S
C2     H1      1.085   .     S
C3     C4      1.401   .     S
C3     H2      1.090   .     S
C4     C5      1.402   .     S
C4     O1      1.364   .     S
C5     C6      1.390   .     S
C5     H3      1.089   .     S
C6     H4      1.091   .     S
C7     C8      1.507   .     S
C7     N1      1.353   .     S
C7     O2      1.252   .     S
C8     H7      1.096   .     S
C8     H8      1.097   .     S
C8     H9      1.100   .     S
N1     H6      1.029   .     S
O1     H5      1.007   .     S
#===END


data_HXACAN26_powder
_audit_creation_date              2012-05-29
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21/A'
```

```
_symmetry_Int_Tables_number       14
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x+1/2,y+1/2,-z
  -x,-y,-z
  x+1/2,-y+1/2,z
_cell_length_a                    12.8856
_cell_length_b                    9.3801
_cell_length_c                    7.1010
_cell_angle_alpha                 90.0000
_cell_angle_beta                  115.7002
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1     C     0.44111   0.35006  -0.14759   0.00000  Uiso   1.00
C2     C     0.51567   0.34905  -0.24418   0.00000  Uiso   1.00
C3     C     0.61093   0.25879  -0.17085   0.00000  Uiso   1.00
C4     C     0.63380   0.16801  -0.00078   0.00000  Uiso   1.00
C5     C     0.55994   0.17013   0.09723   0.00000  Uiso   1.00
C6     C     0.46516   0.26058   0.02478   0.00000  Uiso   1.00
C7     C     0.28233   0.50163  -0.39830   0.00000  Uiso   1.00
C8     C     0.17398   0.57618  -0.41948   0.00000  Uiso   1.00
N1     N     0.34266   0.43758  -0.21130   0.00000  Uiso   1.00
O1     O     0.72742   0.08032   0.06326   0.00000  Uiso   1.00
O2     O     0.31219   0.49756  -0.54401   0.00000  Uiso   1.00
H1     H     0.74879   0.04092   0.20721   0.00000  Uiso   1.00
H2     H     0.49988   0.41887  -0.37568   0.00000  Uiso   1.00
H3     H     0.66924   0.26136  -0.24525   0.00000  Uiso   1.00
H4     H     0.57595   0.09928   0.22844   0.00000  Uiso   1.00
H5     H     0.40813   0.26184   0.10192   0.00000  Uiso   1.00
H6     H     0.14212   0.64609  -0.55615   0.00000  Uiso   1.00
H7     H     0.10720   0.49683  -0.43834   0.00000  Uiso   1.00
H8     H     0.18906   0.63899  -0.27965   0.00000  Uiso   1.00
H9     H     0.30771   0.44477  -0.10583   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
_ccdc_geom_bond_type
C1     C2      1.402   .     S
C1     C6      1.403   .     S
C1     N1      1.411   .     S
```

```
C2     C3      1.393   .     S
C2     H2      1.085   .     S
C3     C4      1.401   .     S
C3     H3      1.090   .     S
C4     C5      1.402   .     S
C4     O1      1.365   .     S
C5     C6      1.390   .     S
C5     H4      1.089   .     S
C6     H5      1.091   .     S
C7     C8      1.510   .     S
C7     N1      1.354   .     S
C7     O2      1.251   .     S
C8     H6      1.093   .     S
C8     H7      1.101   .     S
C8     H8      1.097   .     S
N1     H9      1.030   .     S
O1     H1      1.007   .     S
#===END


data_NAPHTA11_x-ray_H_added
_audit_creation_date              2012-05-17
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21/A'
_symmetry_Int_Tables_number       14
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x+1/2,y+1/2,-z
  -x,-y,-z
  x+1/2,-y+1/2,z
_cell_length_a                    8.2350
_cell_length_b                    6.0030
_cell_length_c                    8.6580
_cell_angle_alpha                 90.0000
_cell_angle_beta                  122.9200
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1     C     0.08740   0.01708   0.32901   0.00000  Uiso   1.00
C2     C     0.11545   0.16019   0.22091   0.00000  Uiso   1.00
C3     C     0.04843   0.10443   0.03628   0.00000  Uiso   1.00
C4     C     0.07447   0.24835  -0.07854   0.00000  Uiso   1.00
C5     C    -0.00842  -0.18889   0.25730   0.00000  Uiso   1.00
```

```
H6     H     0.13920   0.06349   0.47052   0.00000  Uiso   1.00
H7     H     0.18877   0.31952   0.27531   0.00000  Uiso   1.00
H8     H     0.14860   0.40691  -0.02263   0.00000  Uiso   1.00
H9     H    -0.03090  -0.30104   0.34339   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
_ccdc_geom_bond_type
C1     C2      1.379   .     S
C1     C5      1.416   .     S
C1     H6      1.092   .     S
C2     C3      1.419   .     S
C2     H7      1.092   .     S
C3     C4      1.419   .     S
C3     C3      1.435   3     S
C4     H8      1.092   .     S
C4     C5      1.380   3     S
C5     H9      1.092   .     S
C5     C4      1.380   3     S
#===END


data_NAPHTA36_neutron
_audit_creation_date              2012-05-17
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21/A'
_symmetry_Int_Tables_number       14
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x+1/2,y+1/2,-z
  -x,-y,-z
  x+1/2,-y+1/2,z
_cell_length_a                    8.2560
_cell_length_b                    5.9830
_cell_length_c                    8.6770
_cell_angle_alpha                 90.0000
_cell_angle_beta                  122.7290
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1     C     0.08653   0.01746   0.32788   0.00000  Uiso   1.00
```

```
C2     C     0.11573   0.16022   0.22027   0.00000  Uiso   1.00
C3     C     0.04881   0.10440   0.03619   0.00000  Uiso   1.00
C4     C     0.07582   0.24811  -0.07815   0.00000  Uiso   1.00
C5     C    -0.01018  -0.18841   0.25625   0.00000  Uiso   1.00
H1     H     0.13833   0.06388   0.46899   0.00000  Uiso   1.00
H2     H     0.19013   0.31928   0.27453   0.00000  Uiso   1.00
H3     H     0.15024   0.40679  -0.02236   0.00000  Uiso   1.00
H4     H    -0.03364  -0.30034   0.34189   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
_ccdc_geom_bond_type
C1     C2      1.379   .     S
C1     C5      1.416   .     S
C1     H1      1.092   .     S
C2     C3      1.419   .     S
C2     H2      1.092   .     S
C3     C4      1.419   .     S
C3     C3      1.434   3     S
C4     H3      1.092   .     S
C4     C5      1.380   3     S
C5     H4      1.093   .     S
C5     C4      1.380   3     S
#===END


data_NAPHTHA37_power_H_added
_audit_creation_date              2012-05-17
_audit_creation_method            'Materials Studio'
_symmetry_space_group_name_H-M    'P21/C'
_symmetry_Int_Tables_number       14
_symmetry_cell_setting            monoclinic
loop_
_symmetry_equiv_pos_as_xyz
  x,y,z
  -x,y+1/2,-z+1/2
  -x,-y,-z
  x,-y+1/2,z+1/2
_cell_length_a                    8.6869
_cell_length_b                    6.0123
_cell_length_c                    8.2938
_cell_angle_alpha                 90.0000
_cell_angle_beta                  122.5971
_cell_angle_gamma                 90.0000
loop_
_atom_site_label
_atom_site_type_symbol
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
```

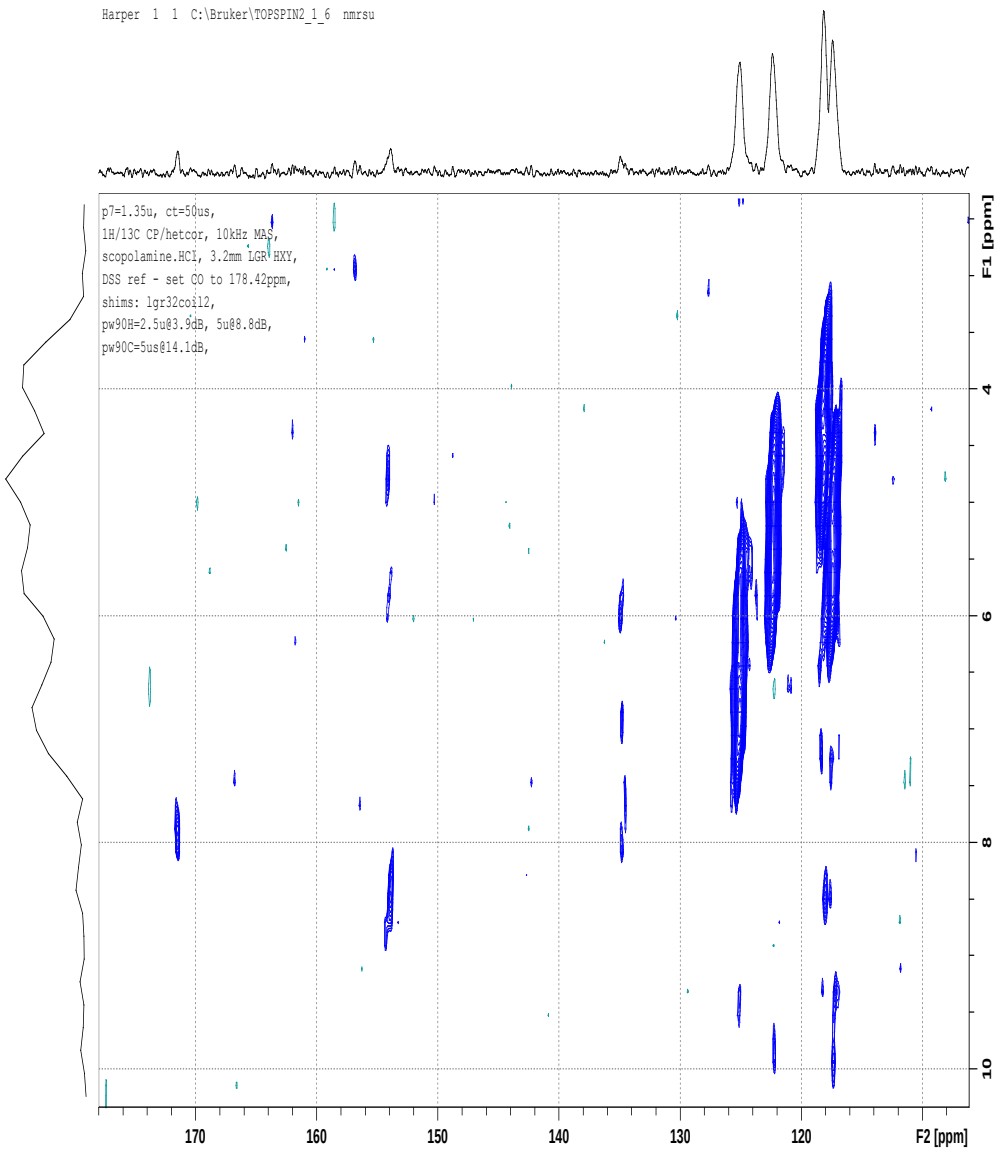
```
_atom_site_U_iso_or_equiv
_atom_site_adp_type
_atom_site_occupancy
C1     C     0.03510   0.10438   0.04808   0.00000  Uiso   1.00
C2     C     0.21833   0.16248   0.11448   0.00000  Uiso   1.00
C3     C     0.32727   0.02184   0.08635   0.00000  Uiso   1.00
C4     C    -0.25785   0.18392   0.00886   0.00000  Uiso   1.00
C5     C    -0.08061   0.24576   0.07410   0.00000  Uiso   1.00
H6     H     0.27102   0.32149   0.18768   0.00000  Uiso   1.00
H7     H     0.46773   0.06977   0.13784   0.00000  Uiso   1.00
H8     H    -0.34463   0.29413   0.03133   0.00000  Uiso   1.00
H9     H    -0.02673   0.40437   0.14753   0.00000  Uiso   1.00
loop_
_geom_bond_atom_site_label_1
_geom_bond_atom_site_label_2
_geom_bond_distance
_geom_bond_site_symmetry_2
_ccdc_geom_bond_type
C1     C2      1.419   .     S
C1     C5      1.419   .     S
C1     C1      1.435   3     S
C2     C3      1.379   .     S
C2     H6      1.092   .     S
C3     H7      1.092   .     S
C3     C4      1.416   3     S
C4     C5      1.380   .     S
C4     H8      1.092   .     S
C4     C3      1.416   3     S
C5     H9      1.092   .     S
#===END
```

^{1}H shift assignmentS, ^{1}H/^{13}C HETCOR spectrum

The following ^{1}H shifts were determined from the ^{1}H/^{13}C HETCOR analysis. Corresponding ^{13}C shifts are included in the manuscript in Table 2-1.

Position	δ ^{1}H (ppm)
C1N – H	7.9
C2 – H	6.6
C3 – H	5.4
C4O – H	8.6
C5 – H	4.4
C6 – H	5.4
C8 – H	-0.4

Figure A-1: An expansion of the $^{1}H/^{13}C$ HETCOR spectrum emphasizing the aromatic, NH (7.9 ppm) and OH (8.6 ppm) protons of acetaminophen. Correlation involving methyl protons have been omitted. Acquisition parameters are included in the manuscript.



APPENDIX B: SUPPLEMENTARY MATERIAL FOR CHAPTER 3

Supporting Information for

Establishing Accurate High-Resolution Crystal Structures in the Absence of Diffraction Data and Single-Crystals – a Solid-State NMR Approach.

Table of Contents.

Refined neutron diffraction coordinates.

The following tables detail the neutron diffraction coordinates refined under lattice constraints (PBE/ultrafine) for the four carbohydrate studied. In all cases the unit cell parameters reported in the original neutron diffraction study were not altered since it is known that the refinement methods used systematically overestimates these parameters.[1]

Table B-1: Refined Neutron Coordinates for Methyl α-D-galactopyranoside under PBE/Ultrafine lattice constraints.

Unit cell parameters

a	7.779	α	90
b	8.535	β	90
c	13.131	γ	90

Space group: $P2_12_12_1$.

27

H	1.5223	4.3138	11.752
H	0.7841	6.7097	13.533
H	0.3228	6.2409	10.55
H	1.9416	7.8706	10.022
H	2.7075	5.5616	10.285
H	4.4222	7.2819	9.7831
H	4.7641	7.4958	11.502
H	2.4919	2.6067	14.654
H	3.5442	3.2136	13.341
H	1.9748	2.4298	12.953
H	-0.954	5.4136	13.732
H	-0.481	8.2829	12.317
H	2.8335	9.1812	11.358
H	5.924	5.6743	11.5
C	1.7302	5.0224	12.58
C	0.6031	6.0631	12.659
C	0.6045	6.9098	11.384
C	1.9968	7.454	11.042
C	3.0084	6.3006	11.053
C	4.4249	6.7653	10.756
C	2.5026	3.0808	13.668
O	1.8305	4.3418	13.797
O	-0.669	5.425	12.773
O	-0.382	7.9326	11.41
O	2.3637	8.5124	11.936
O	2.9883	5.6594	12.339
O	5.3349	5.6691	10.7

Table B-2: Refined Neutron Coordinates for Methyl α-D-glucopyranoside under PBE/Ultrafine lattice constraints.

Unit cell parameters

a	11.311	α	90
b	14.781	β	90
c	5.281	γ	90

Space group: $P2_12_12_1$.

27

H	-3.914	6.0816	2.0207
H	-6.319	5.9957	1.5381
H	-6.789	6.9507	4.4062
H	-7.456	4.2715	3.1075
H	-5.269	4.8919	5.1711
H	-5.719	2.3353	3.5663
H	-6.382	2.6132	5.202
H	-2.788	5.5441	4.8323
H	-2.572	7.3061	5.0689
H	-2.164	6.5608	3.4944
H	-5.851	8.5465	2.8031
H	-8.87	7.2077	3.7388
H	-8.696	5.0158	5.0586
H	-3.662	2.2897	4.4043
C	-4.636	6.0373	2.8556
C	-6.007	6.5704	2.4242
C	-7.041	6.334	3.5284
C	-7.041	4.8578	3.949
C	-5.619	4.3736	4.2611
C	-5.57	2.878	4.5142
C	-2.859	6.5311	4.35
O	-4.192	6.8294	3.9328
O	-5.926	7.9293	2.0118
O	-8.327	6.7224	3.0433
O	-7.796	4.6375	5.1379
O	-4.723	4.654	3.1728
O	-4.341	2.4678	5.1219

Table B-3: Refined Neutron Coordinates for Methyl α-D-mannopyranoside under PBE/Ultrafine lattice constraints.

Unit cell parameters

a	9.429	α	90
b	9.315	β	90
c	10.055	γ	90

Space group: $P2_12_12_1$.

27

H	2.3857	0.7006	3.4893
H	4.3673	-0.007	1.9994
H	5.9499	-0.631	3.7355
H	6.0191	2.3125	4.5671
H	4.8517	-0.114	6.0401
H	3.9296	1.4921	7.5963
H	5.7063	1.5494	7.5691
H	1.3958	-1.0000	4.7738
H	2.3386	-2.515	4.9839
H	2.6211	-1.126	6.0771
H	5.1096	2.3378	1.972
H	6.9075	0.1838	1.8392
H	7.8013	1.6513	5.4666
H	3.8501	3.5001	6.8578
C	3.375	0.39	3.8688
C	4.4842	0.6802	2.8578
C	5.8573	0.442	3.502
C	5.9923	1.2372	4.7957
C	4.8075	0.9451	5.7266
C	4.8003	1.7921	6.9933
C	2.3745	-1.423	5.0494
O	3.3921	-0.996	4.1332
O	4.3093	2.0315	2.4555
O	6.8957	0.8332	2.5964
O	7.2039	0.8682	5.4755
O	3.5525	1.175	5.0464
O	4.7833	3.1921	6.7408

Table B-4: Refined Neutron Coordinates for Methyl β-D-xylopyranoside under PBE/Ultrafine lattice constraints.

Unit cell parameters

a	7.877	α	90
b	6.933	β	113.38
c	7.748	γ	90

Space group: $P2_1$.

23

H	-2.941	1.9088	2.5609
H	-2.971	2.9297	-0.35
H	-0.685	2.5478	1.6555
H	-1.899	5.1675	0.6864
H	-2.318	3.8935	3.4457
H	-3.028	5.4139	2.8159
H	-5.89	2.3182	2.84
H	-6.187	0.6938	2.145
H	-4.822	0.9573	3.2988
H	-2.781	0.7623	-0.614
H	0.4636	3.8019	0.0025
H	0.0084	5.6789	1.4796
C	-3.427	2.3518	1.6677
C	-2.47	2.3864	0.4674
C	-1.216	3.1502	0.892
C	-1.558	4.5166	1.508
C	-2.677	4.4122	2.5379
C	-5.415	1.3932	2.4808
O	-4.582	1.6595	1.3394
O	-2.109	1.0689	0.0384
O	-0.373	3.3306	-0.249
O	-0.413	5.0724	2.1448
O	-3.802	3.7079	1.9944

References and Footnotes.

[1]. Ashbrook, S. E.; Cutajar, M.; Pickard, C. J.; Walton, R. I.; Wimperis, S. *Phys. Chem. Chem. Phys.*, **2008**, *10*, 5754-5764.

APPENDIX C: SUPPLEMENTARY MATERIAL FOR CHAPTER 4

Contents.

Coordinates for DFT refined structures.

Table C-1: Coordinates for DFT refined Acetaminophen (CSD refcode: HXACAN26)

	20		
H	9.0105	0.3838	1.3259
H	7.5981	3.929	-2.404
H	9.3787	2.4516	-1.569
H	6.7181	0.9312	1.4617
H	4.9452	2.4561	0.6522
H	3.544	6.0604	-3.559
H	2.7312	4.6604	-2.805
H	3.2973	5.9938	-1.789
H	4.2909	4.172	-0.677
C	6.1384	3.2836	-0.944
C	7.3967	3.2741	-1.562
C	8.3984	2.4275	-1.093
C	8.1693	1.576	-0.005
C	6.9158	1.5958	0.6222
C	5.9176	2.4443	0.1585
C	4.8645	4.7054	-2.549
C	3.5336	5.4046	-2.684
N	5.0661	4.1045	-1.352
O	9.1784	0.7534	0.4048
O	5.6979	4.6672	-3.481

Table C-2: Coordinates for DFT refined Cimetidine, form A (CSD refcode: CIMETD03)

33			
H	4.8102	5.9415	-1.836
H	2.4169	5.2813	-1.258
H	0.3942	0.6342	2.5279
H	1.6812	2.5986	-0.221
H	7.3297	4.9851	-0.917
H	7.0988	4.0108	-2.386
H	7.3782	3.2158	-0.831
H	4.4247	0.9119	-0.373
H	6.0607	1.5205	-0.109
H	2.8109	2.8641	1.8445
H	2.9763	1.8254	3.2832
H	2.4817	-0.188	1.9968
H	2.4984	0.6989	0.4575
H	-0.712	3.8295	-0.989
H	0.6168	3.3195	-2.056
H	-0.763	2.2292	-1.776
C	3.284	4.6455	-1.09
C	4.6022	3.0356	-0.488
C	5.4096	4.0141	-1.047
C	6.873	4.0643	-1.309
C	4.9999	1.7279	0.087
C	3.1048	1.8637	2.1916
C	2.2551	0.779	1.5258
C	0.0796	1.7024	0.7816
C	-0.102	2.9701	-1.307
C	-1.937	0.8955	1.5968
N	4.5451	5.023	-1.422
N	3.276	3.4408	-0.531
N	0.8254	1.0027	1.6562
N	0.655	2.3876	-0.214
N	-1.277	1.7592	0.8732
N	-2.655	0.1602	2.1887
S	4.8774	1.6143	1.9186

Table C-3: Coordinates for DFT refined Glycine, γ-phase (CSD refcode: GLYCIN18)

	10		
H	-1.758	-4.647	-0.459
H	-0.069	-4.942	-0.492
H	-2.177	-7.041	-0.448
H	-0.432	-7.29	-0.654
H	-0.988	-5.427	0.8648
C	-1.263	-6.607	-0.871
C	-1.382	-6.442	-2.384
N	-0.992	-5.324	-0.194
O	-1.879	-7.41	-3.021
O	-0.936	-5.371	-2.906

Table C-4: Coordinates for DFT refined Thymine (refcode: THYMIN01)

	15		
H	7.0621	1.9962	2.275
H	7.0501	-2.075	2.3644
H	9.0618	2.1443	3.5438
H	10.91	1.026	4.7448
H	10.362	-0.468	5.5535
H	11.34	-0.567	4.0808
C	6.9074	-0.042	2.2355
C	8.701	-1.274	3.3998
C	9.2981	0.004	3.7577
C	8.6862	1.1417	3.3282
C	10.545	0.0057	4.5779
N	7.5293	1.1176	2.5857
N	7.5129	-1.189	2.6565
O	5.8401	-0.051	1.5692
O	9.1511	-2.386	3.7061

Table C-5: Coordinates for DFT refined Histidine hydrochloride hydrate (CSD refcode: HISTCM01)

	25		
H	4.3146	-7.995	1.2079
H	5.5748	-4.436	2.7107
H	5.528	-9.086	0.6718
H	4.7496	-9.336	2.1653
H	6.12	-5.898	-1.305
H	4.4409	-2.893	0.9675
H	6.9116	-8.58	2.5694
H	7.2023	-7.405	0.3795
H	7.8336	-6.567	1.7873
H	4.8955	-3.652	-1.39
H	5.913	-1.12	5.4375
H	7.2452	-1.309	6.2549
C	5.6189	-7.159	3.4937
C	6.1744	-7.839	2.2287
C	6.9016	-6.864	1.2878
C	6.1308	-5.643	0.8967
C	5.197	-4.18	-0.487
C	5.5846	-4.624	1.6456
N	5.1154	-8.599	1.5207
N	5.8748	-5.328	-0.425
N	5.0089	-3.739	0.7623
O	4.401	-7.269	3.7569
O	6.4931	-6.52	4.1728
O	6.4076	-0.781	6.225
Cl	4.9464	-1.97	3.7122

Changes in atomic positions from DFT refinement.

The DFT refinement of the original diffraction coordinates resulted in new atomic coordinates for each of the 5 structures evaluated. The root-mean squared deviation (RMSD) in atom positions from the refinements (original vs. refined) are given below and includes values for all atoms, non-hydrogen only and hydrogen atoms.

Table C-6: The RMS deviation in atomic positions resulting from the lattice-including DFT secondary refinements of diffraction coordinates.

	RMSD (Å)		
Compound	All-atoms	Non-hydrogen only	Hydrogen only
Acetaminophen	0.179	0.053	0.265
Cimetidine, form A	0.520	0.070	0.762
Glycine, γ-phase	0.586	0.018	0.868
Histidine HCl H2O	0.128	0.054	0.181
Thymine	0.121	0.153	0.203